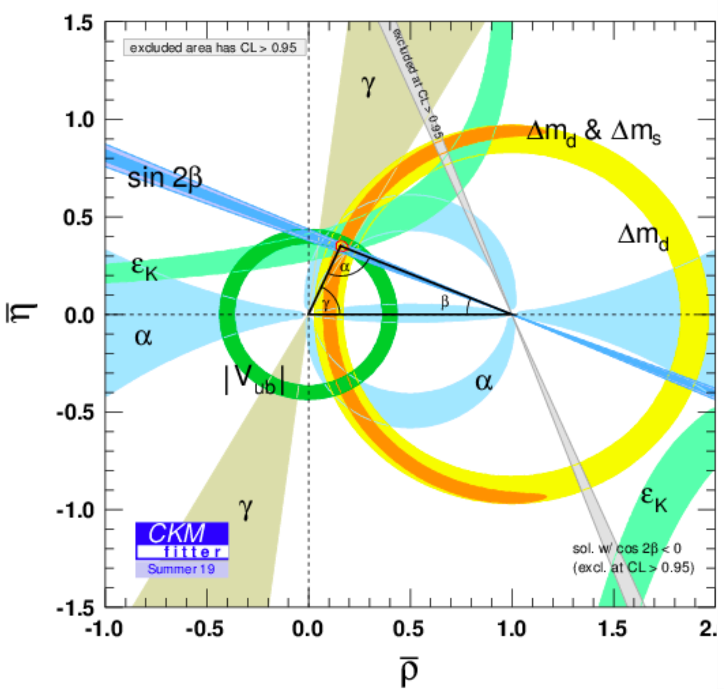
Systematic uncertainties in flavour physics analyses

PHYSTAT 2021

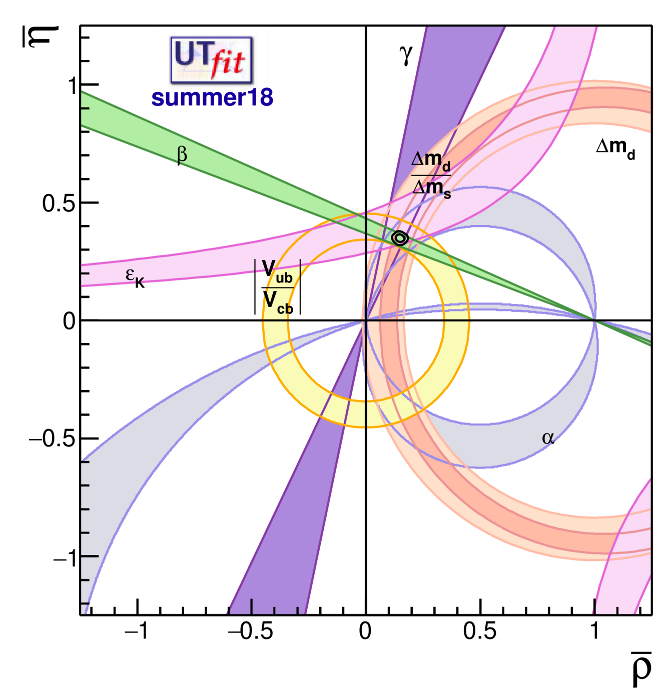
T. Blake
on behalf of the LHCb collaboration

Introduction

- "Flavour physics" covers a broad range of different types of measurements:
  - Branching fraction (rate) measurements and *CP* asymmetries, time-dependent measurements of oscillation parameters, angular and amplitude analyses.
- Often involve complex multidimensional likelihood fits.
- Measurements are often used as input to analyses testing SM consistency, *e.g.* to analyses of CKM parameters or Wilson coefficients.



[EPJC 41 (2005) 1-13]



[JHEP 07 (2001) 013]

Introduction

- Typically consider systematic uncertainties:
  - From statistical uncertainties on external inputs.
  - Due to biasing effects in the measurement.

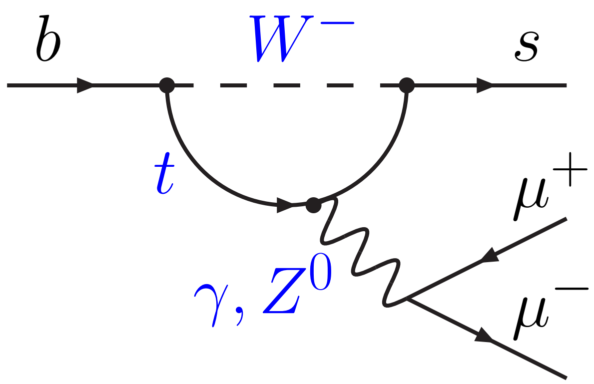
    e.g. due to momentum or length scale uncertainties, or analysis choices.
- No one size fits all approach to treating systematic uncertainties.

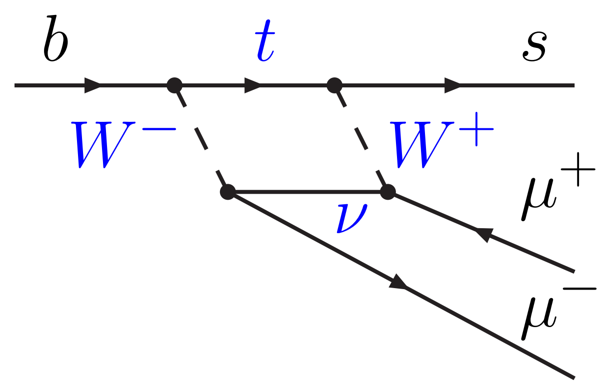
Use LHCb's analysis of the angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$ decays to illustrate some of the issues involved

[LHCb, PRL 125 (2020) 011802]

Why study $B^0 \to K^{*0}\mu^+\mu^-$ decays?

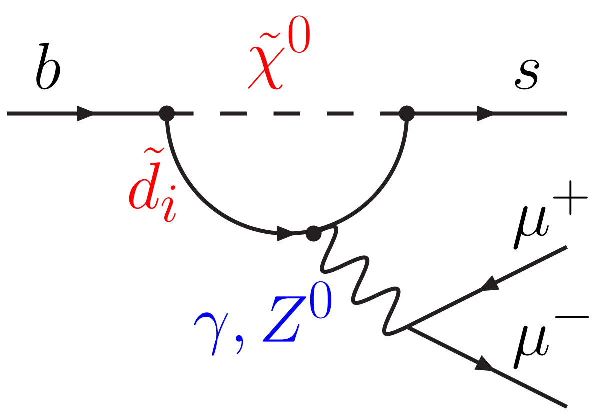
- Flavour changing neutral current transitions only occur at loop order (and beyond) in the SM.

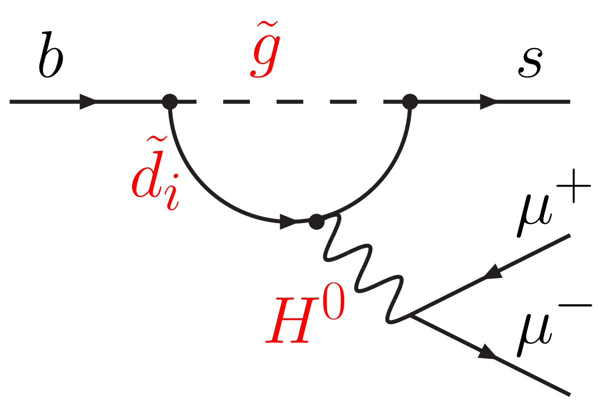


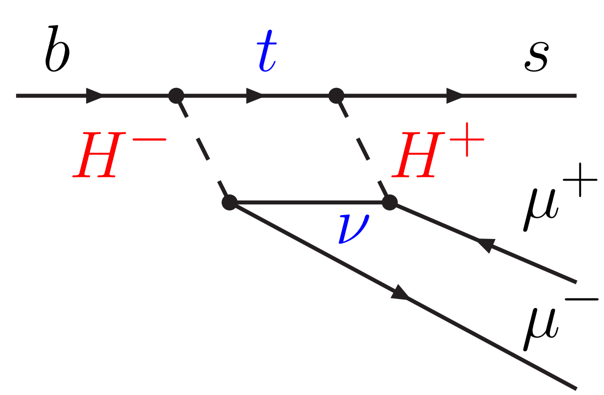


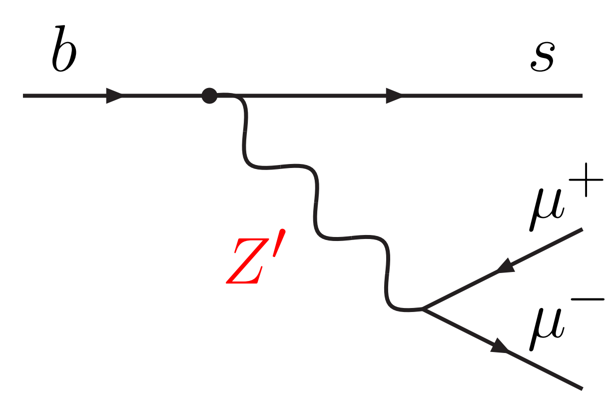
SM diagrams involve the charged current interaction.

- New particles can also contribute:





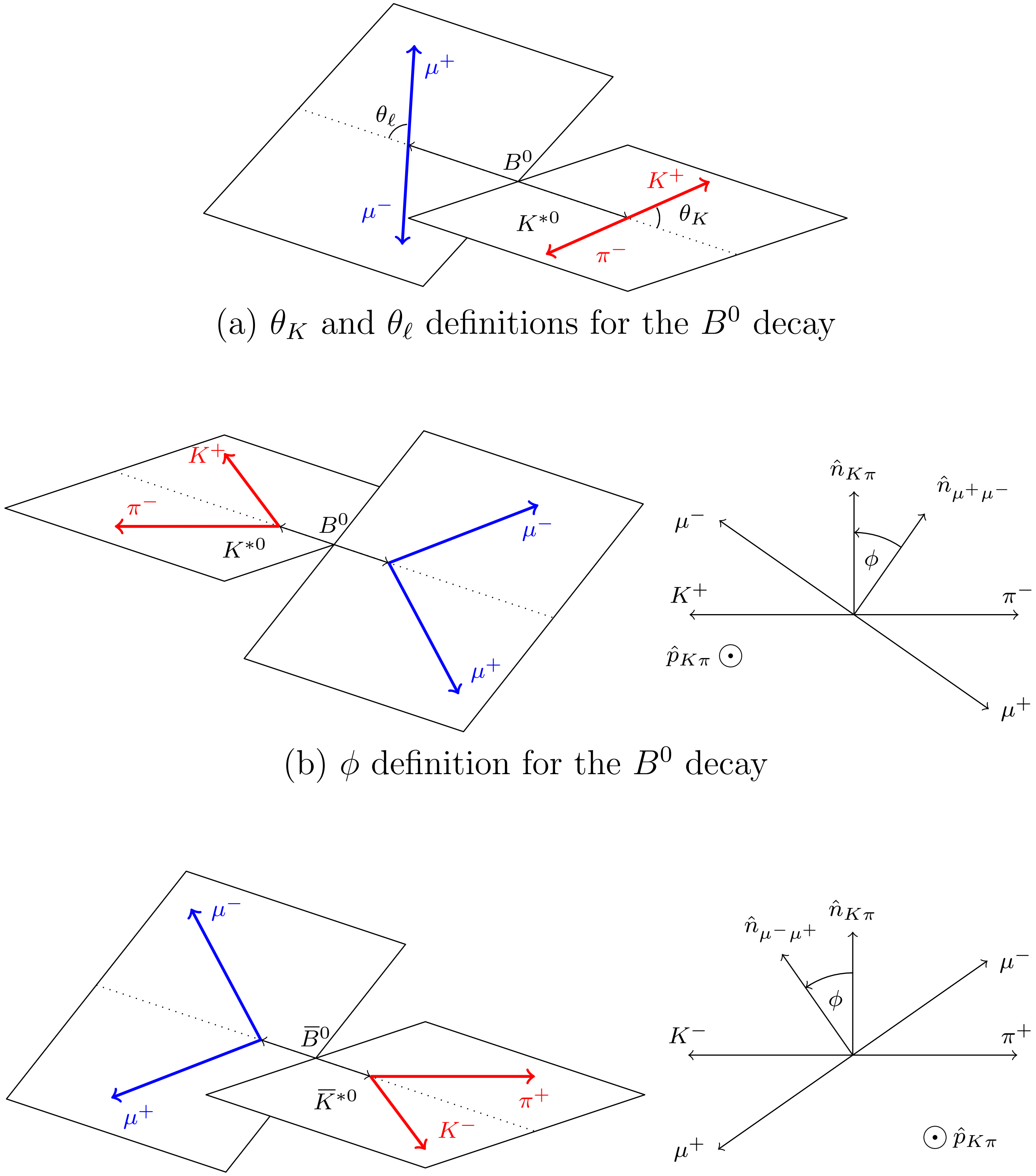




Enhancing/suppressing decay rates or **modifying the angular distribution** of the final-state particles.

Angular variables

- The $B^0 \to K^{*0}\mu^+\mu^-$ decay can be described by:
  - One angle in the $\mu^+\mu^-$ rest-frame.
  - One angle in the K^{*0} rest-frame.
  - The angle between the $\mu^+\mu^-$ and the K^{*0} decay planes.
  - The dimuon invariant mass squared, q^2.



(a) θ_K and θ_ℓ definitions for the B^0 decay

(b) ϕ definition for the B^0 decay

(c) ϕ definition for the $\overline{B}^0$ decay

$B^0 \to K^{*0}\mu^+\mu^-$ angular observables

Complex angular distribution:

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^3(\Gamma+\bar{\Gamma})}{\mathrm{d}\vec{\Omega}}\Big|_{\mathrm{P}} = \frac{9}{32\pi}\Big[\tfrac{3}{4}(1-\boxed{F_{\mathrm{L}}})\sin^2\theta_K + \boxed{F_{\mathrm{L}}}\cos^2\theta_K +$$

$$+\tfrac{1}{4}(1-\boxed{F_{\mathrm{L}}})\sin^2\theta_K\cos 2\theta_l$$

$$-\boxed{F_{\mathrm{L}}}\cos^2\theta_K\cos 2\theta_l + S_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi$$

$$+S_4\sin 2\theta_K\sin 2\theta_l\cos\phi + \boxed{S_5}\sin 2\theta_K\sin\theta_l\cos\phi$$

$$+\tfrac{4}{3}\boxed{A_{\mathrm{FB}}}\sin^2\theta_K\cos\theta_l + S_7\sin 2\theta_K\sin\theta_l\sin\phi$$

$$+S_8\sin 2\theta_K\sin 2\theta_l\sin\phi + S_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\Big]$$

fraction of longitudinal polarisation of the K*

forward-backward asymmetry of the dilepton system

The observables depend on form-factors for the $B^0 \to K^{*0}$ transition plus the underlying short distance physics (Wilson coefficients).

Experiments can reduce the complexity by folding the angular distribution, see **[LHCb, PRL 111 (2013) 191801]**

$B^0 \to K^{*0}\mu^+\mu^-$ angular observables

Complex angular distribution:

$$
\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^3(\Gamma+\bar{\Gamma})}{\mathrm{d}\vec{\Omega}}\bigg|_{\mathrm{P}} = \frac{9}{32\pi}\Big[\tfrac{3}{4}(1-\boxed{F_{\mathrm{L}}})\sin^2\theta_K + \boxed{F_{\mathrm{L}}}\cos^2\theta_K +
$$

$$
+\tfrac{1}{4}(1-\boxed{F_{\mathrm{L}}})\sin^2\theta_K\cos 2\theta_l
$$

$$
-\boxed{F_{\mathrm{L}}}\cos^2\theta_K\cos 2\theta_l + S_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi
$$

$$
+S_4\sin 2\theta_K\sin 2\theta_l\cos\phi + \boxed{S_5}\sin 2\theta_K\sin\theta_l\cos\phi
$$

$$
+\tfrac{4}{3}\boxed{A_{\mathrm{FB}}}\sin^2\theta_K\cos\theta_l + S_7\sin 2\theta_K\sin\theta_l\sin\phi
$$

$$
+S_8\sin 2\theta_K\sin 2\theta_l\sin\phi + S_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\Big]
$$

fraction of longitudinal polarisation of the K*

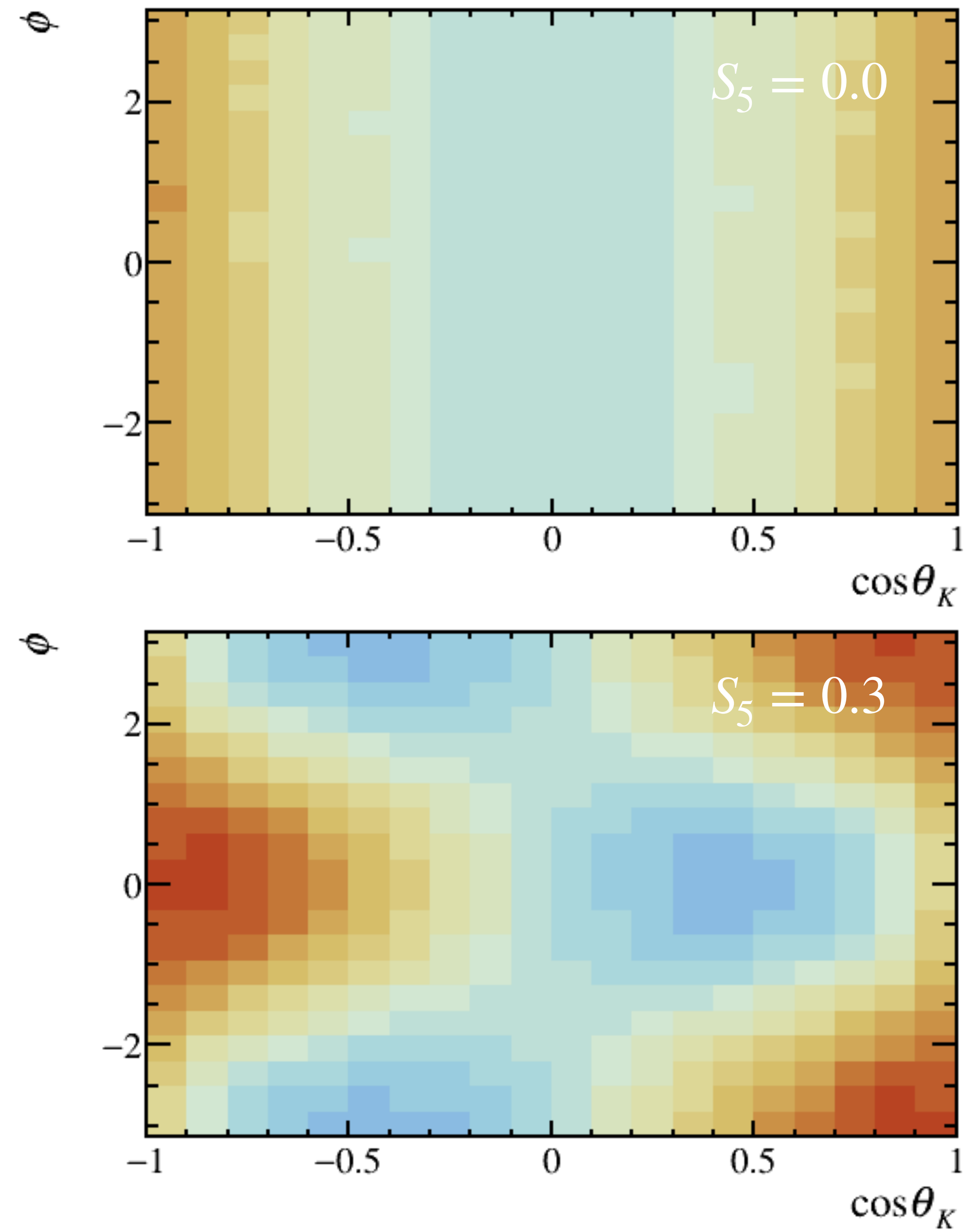
forward-backward asymmetry of the dilepton system

Parameters of interest are q^2 dependent. Measurements correspond to a rate average over a q^2 bin, *i.e.*

$$
S_i = \int \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2} S_i(q^2)\mathrm{d}q^2 \Big/ \int \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2}\mathrm{d}q^2
$$

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

- The functional form of the angular distribution is fixed for a given K^{*0} state (for a given J^P).
- Look for variations in the value of the coefficients, *e.g.* varying the value of the coefficient S_5 changes the distribution of the decay in $(\cos\theta_K, \phi)$.



Maximum likelihood fit

- Angular observables determined using an unbinned maximum likelihood fit to 5 variables:

$$\vec{\Omega} = (\cos\theta_l, \cos\theta_K, \phi, m(K^\pm\pi^\mp\mu^+\mu^-), m(K^\pm\pi^\mp))$$

Angular variables

Mass of B candidate (used to separate signal from background)

Mass of K^* candidate (used to separate S- from P-wave states)

- Eight parameters of interest (F_{L}, A_{FB} and S_i).
- Nuisance parameters associated with experimental backgrounds, S-wave contribution (and S-P interference).
- Simultaneous fit to two data taking periods (Run 1 and 2016 dataset). Typically have $\mathcal{O}(200)$ decays per period in each q^2 bin.

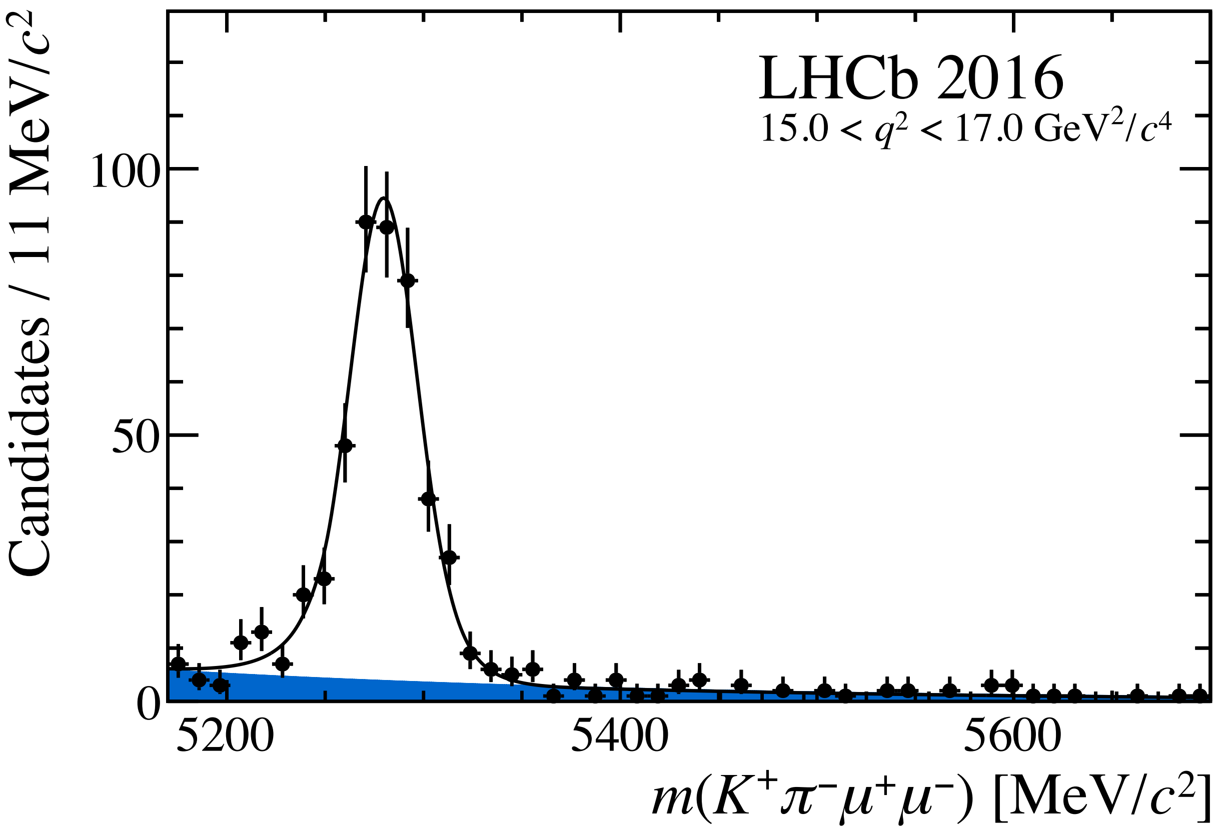
Example fit

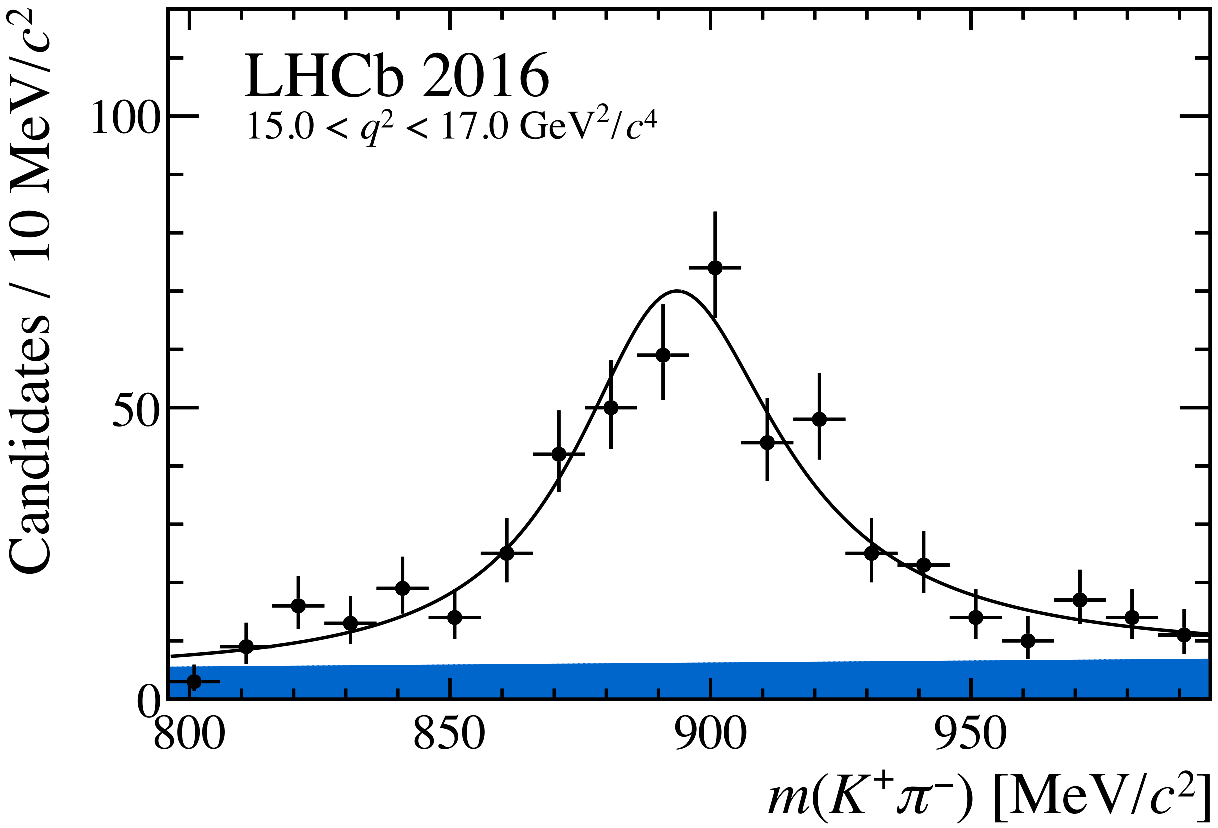
- Projecting the fit result onto the 5 variables:

Signal

Background

[LHCb, PRL 125 (2020) 011802]





Signal "peaks" at the known B^0 mass with width given by experimental resolution.

Signal has a characteristic Breit-Wigner line shape associated with the $K^*(892)^0$.

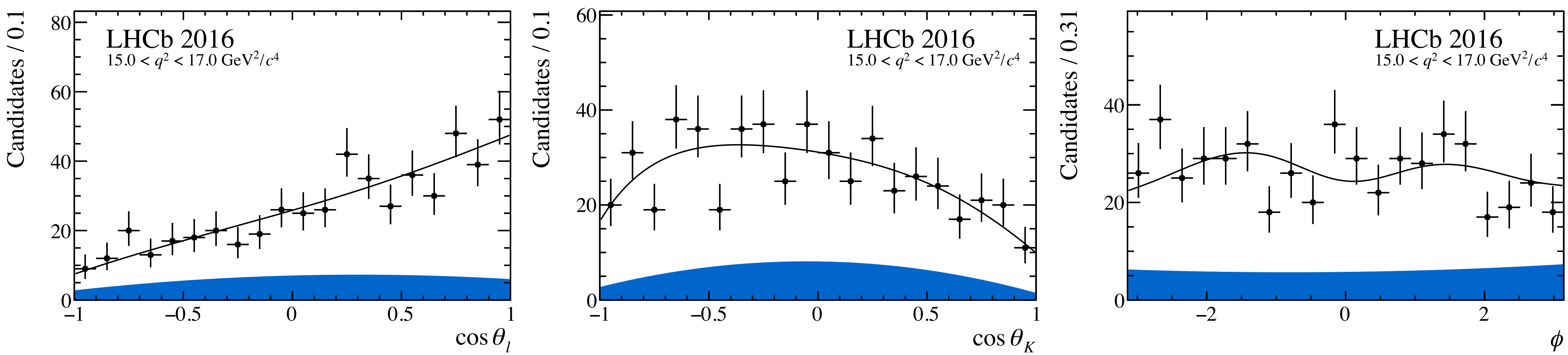
Example fit

- Projecting the fit result onto the 5 variables:

Signal

Background

[LHCb, PRL 125 (2020) 011802]



Large asymmetry generated by non-zero A_{FB}

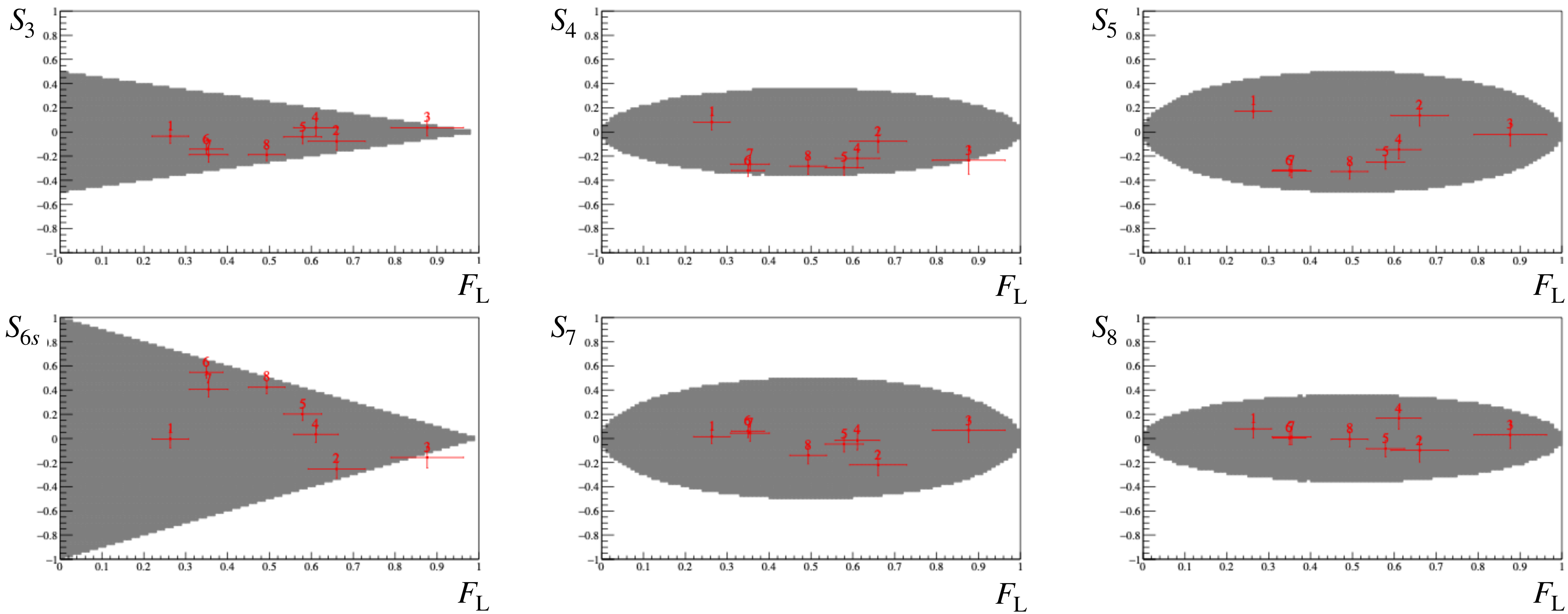
Asymmetry generated by angular acceptance (non-uniform detector efficiency with momentum)

$\cos 2\phi$ dependence generated by non-zero S_3 (and non-uniform angular acceptance)

Statistical uncertainties

- Estimate the covariance matrix for each q^2 bin using the Hessian matrix from the likelihood fit.
- The coverage of the statistical uncertainty is checked using pseudoexperiments. These are generated from:
  - The Standard Model point.
  - The closest physical point to the best-fit point in data.
- A small bias is observed in some observables.
  1. Due to small size of F_S, which is typically biased to larger values.
  2. Due to the proximity of unphysical regions of parameter space.
     (where the PDF becomes negative for some region of the angular variables)

Allowed parameter space



- Several measurements are close to the edge of the allowed range.

PDF remains positive definite

Measurement (with Run 1 data)
[LHCb, JHEP 02 (2016) 104]

Systematic uncertainties

- Consider a range of sources that could bias the angular observables or the estimate of the signal yield.
- Rare process, systematic uncertainties are small compared to the statistical uncertainty (typically 0.04 on the S_i).

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	< 0.03	< 0.09
$m(K^+\pi^-)$ model	< 0.01	< 0.01	< 0.02
Background model	< 0.01	< 0.01	< 0.03
Peaking backgrounds	< 0.02	< 0.02	< 0.03
$m(K^+\pi^-\mu^+\mu^-)$ model	< 0.01	< 0.01	< 0.02
$K^+\mu^+\mu^-$ veto	< 0.01	< 0.01	< 0.01
Trigger	< 0.01	< 0.01	< 0.01
Bias correction	< 0.02	< 0.02	< 0.04

Supplemental material of [PRL 125 (2020) 011802]

- Correlations between systematic uncertainties in different q^2 bins are neglected (no significant correlation is seen in pseudoexperiments).

Remarks

- Two classes of systematic uncertainty:

  1. Uncertainties due to analysis choices, which are typically binary variations (chosen as an example of other possible options).

     If there are multiple choices of model for the signal, *e.g.* choice of resonances in an amplitude fit, it is usual to quote fit results for every model that has a similar fit quality.

  2. Statistical uncertainties on external inputs.
     e.g. due to limited MC sample sizes or calibration sample sizes.

Remarks

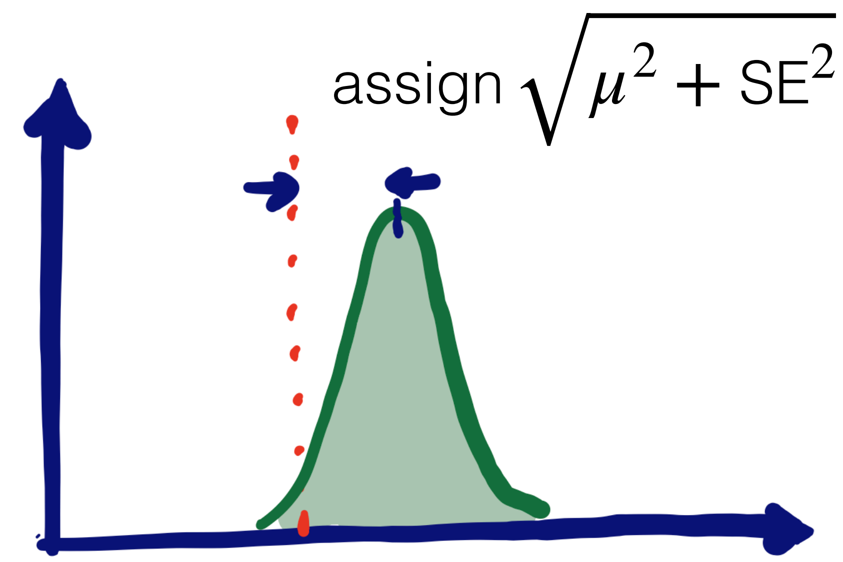
- Two classes of systematic uncertainty:
  1. Uncertainties due to analysis choices, which are typically binary variations (chosen as an example of other possible options).
  2. Statistical uncertainties on external inputs.
- Possible issues?
  - Not clear that these uncertainties can be treated as normally distributed or how they can be combined with the statistical uncertainty in a meaningful way.
  - No concept of a 68% interval for class (1).

Systematic approaches

- Two different approaches to estimating systematic uncertainties:

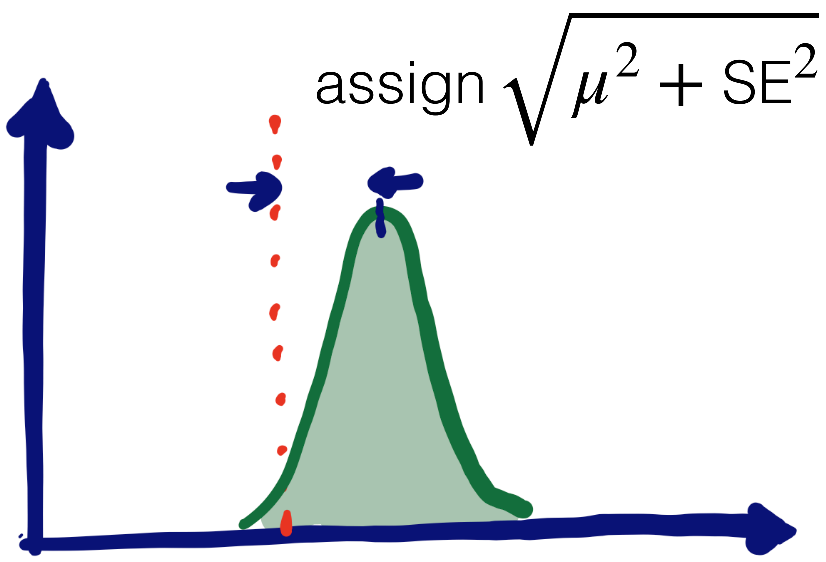
1. Generate a large number of pesudoexperiments from a varied model and determine observables using the nominal model.

   i.e. estimate of expected bias

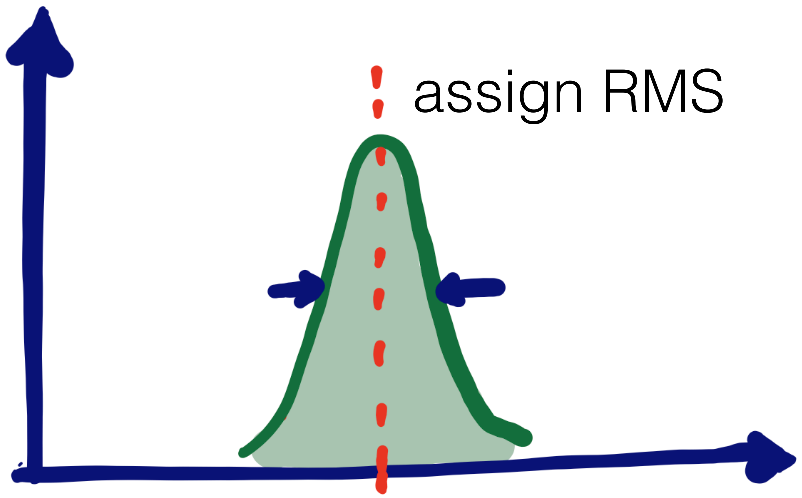


Systematic aproaches

- Two different approaches to estimating systematic uncertainties:



1. Generate a large number of pesudoexperiments from a <u>varied model</u> and determine observables using the <u>nominal model</u>.



2. Repeat the determination of the observables in data using a different set of assumptions.

Acceptance statistical uncertainty

- Impact of the detector on the signal angular distribution is parameterised by polynomials in 4-dimensions:

$$\epsilon(\cos\theta_l, \cos\theta_K, \phi, q^2) = \sum_{jklm} c_{jklm} P_j(\theta_l) P_k(\theta_K) P_l(\phi) P_m(q^2)$$

Legendre polynomials

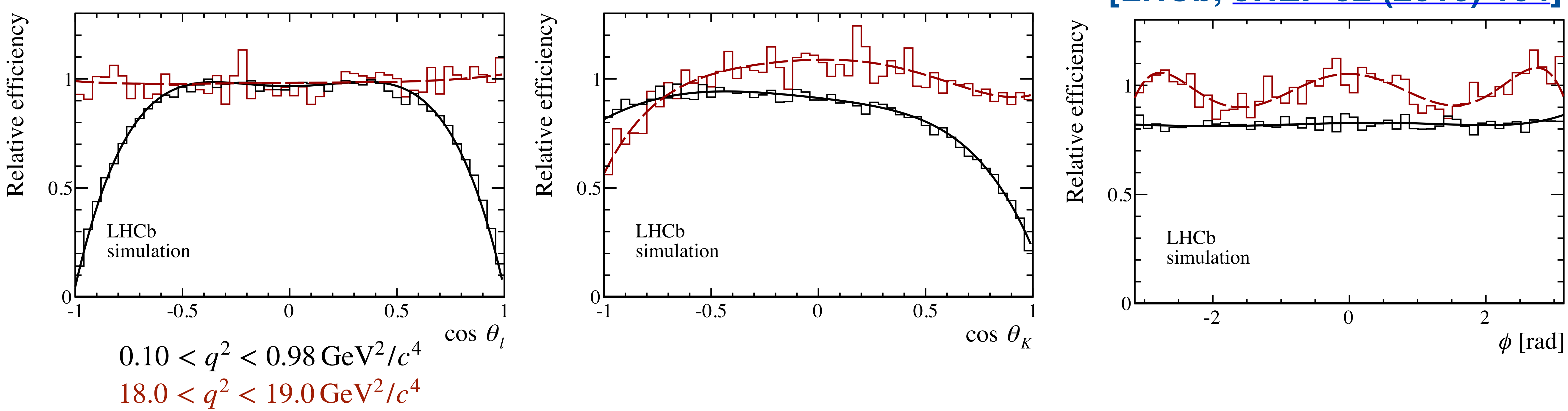
- Efficiency shape is fixed from simulation in the fit to data (due to the large number of coefficients, c_{jklm}).

- Evaluate systematic by:
    - Bootstrapping (sampling with replacement) the simulated samples and deriving a new efficiency model.
    [B. Efron, Annals Statist. 7 (1979) 1, 1-26]
    - Refitting the data with the new efficiency model.

Acceptance variation with q^2

- Angular efficiency shape varies strongly with q^2:

[LHCb, JHEP 02 (2016) 104]



  - Shape is primarily due to the LHCb angular coverage and momentum dependent reconstruction efficiencies.

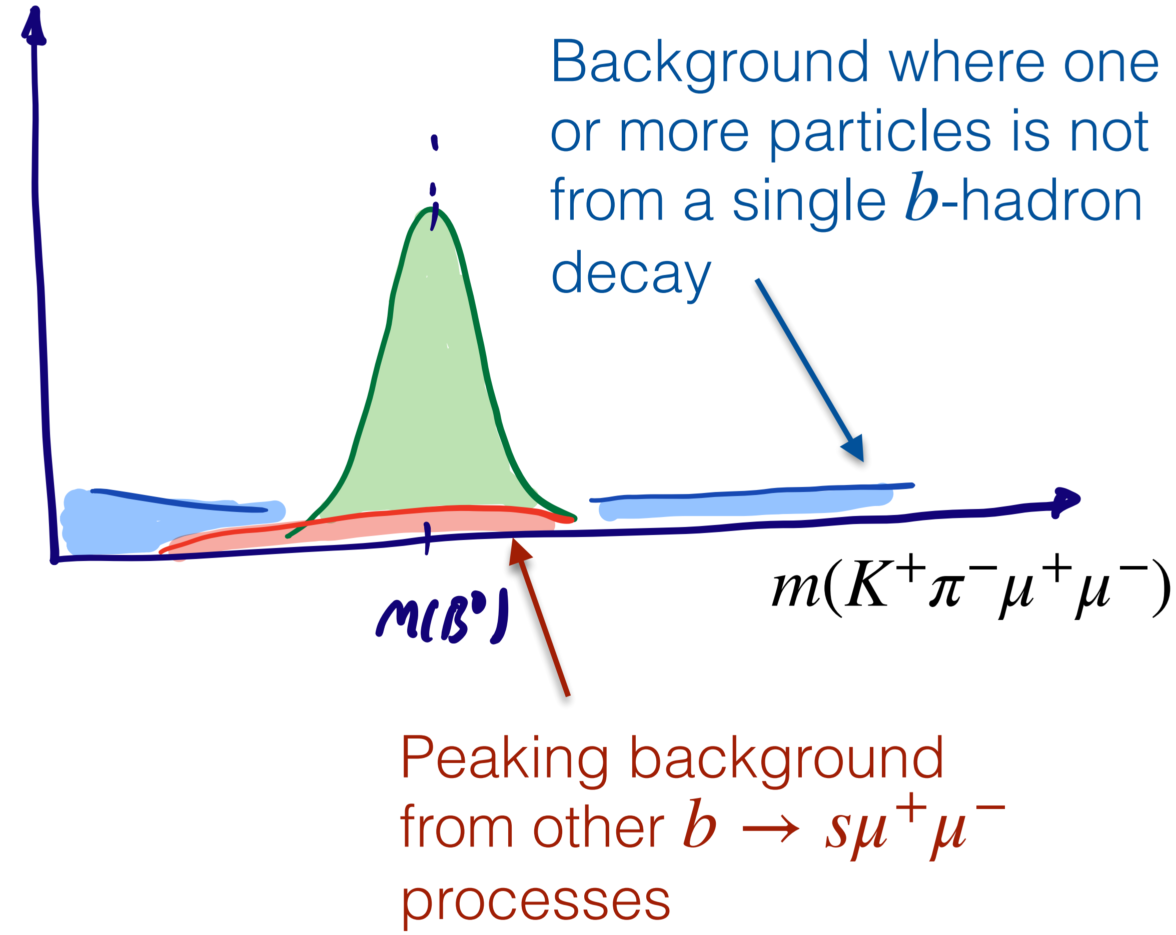
- Shape is fixed in the fits to each bin at the q^2 value of the bin centre.

Acceptance variation with q^2

- The acceptance variation over the q^2 bin is neglected.
- Consider two systematic variations:
  - Point halfway between the bin centre and the upper bin edge.
  - Point halfway between the lower bin edge and the bin centre.
- Generate pseudoexperiments based on the varied acceptance model and fit back with the nominal model.
- Note, we also measure observables in wide q^2 bins by weighting candidates in the likelihood fit (associated with its own problems).

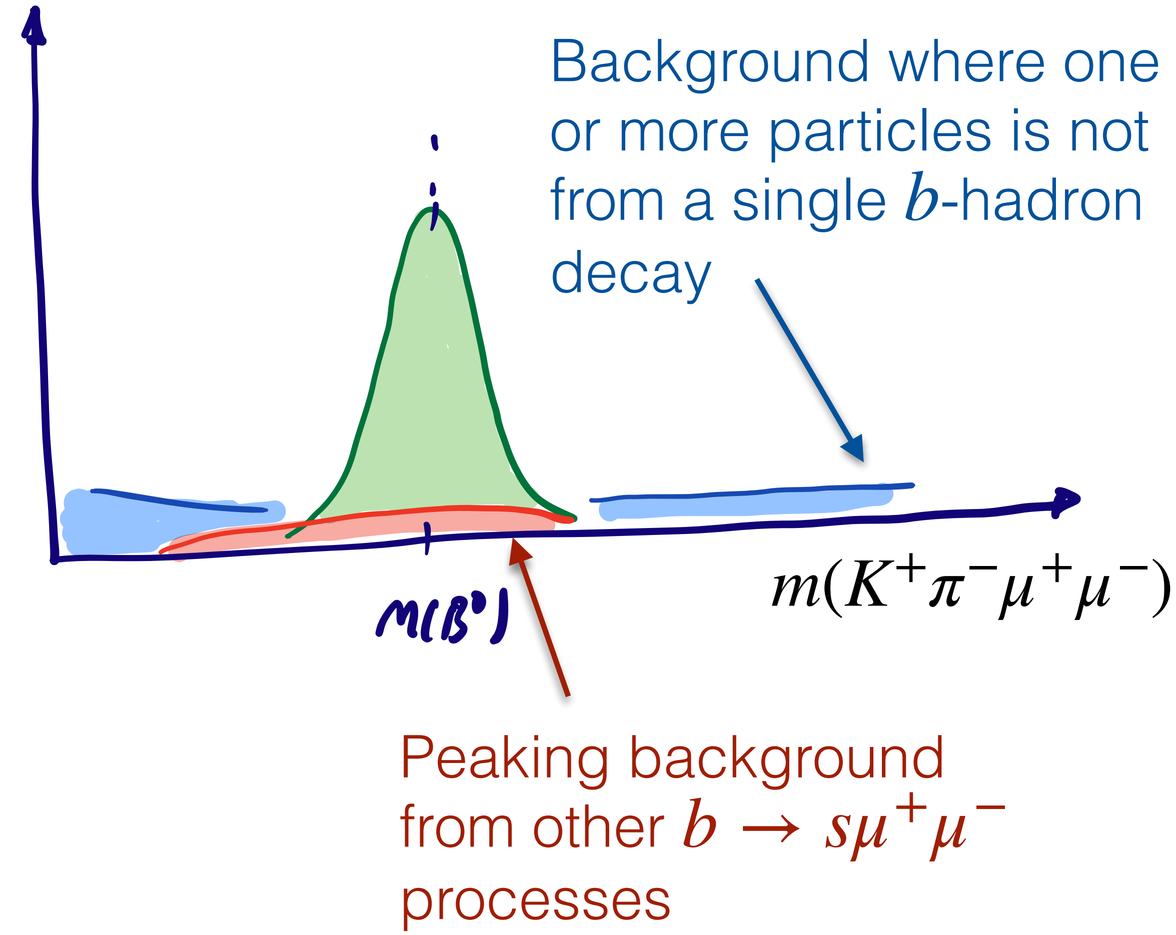
"Peaking" backgrounds

- There are a number of backgrounds from other $b \to s\mu^+\mu^-$ processes that could mimic the signal ($\Lambda_b \to pK^-\mu^+\mu^-,\ B^0_s \to K^+K^-\mu^+\mu^-, \ldots$).
    - Same topology as the signal but with one or more hadrons incorrectly identified.
    - Suppressed to $< 1\,\%$ of the signal yield using information from LHCb's RICH detectors.
    - Residual backgrounds are small and neglected in the likelihood fit.



“Peaking” backgrounds

- Estimate systematic uncertainty from neglecting the background contributions by:
  - Generating pseudoexperiments where a background contribution is injected.
  - Fitting the pseudodata, neglecting the background contributions as is done in data.

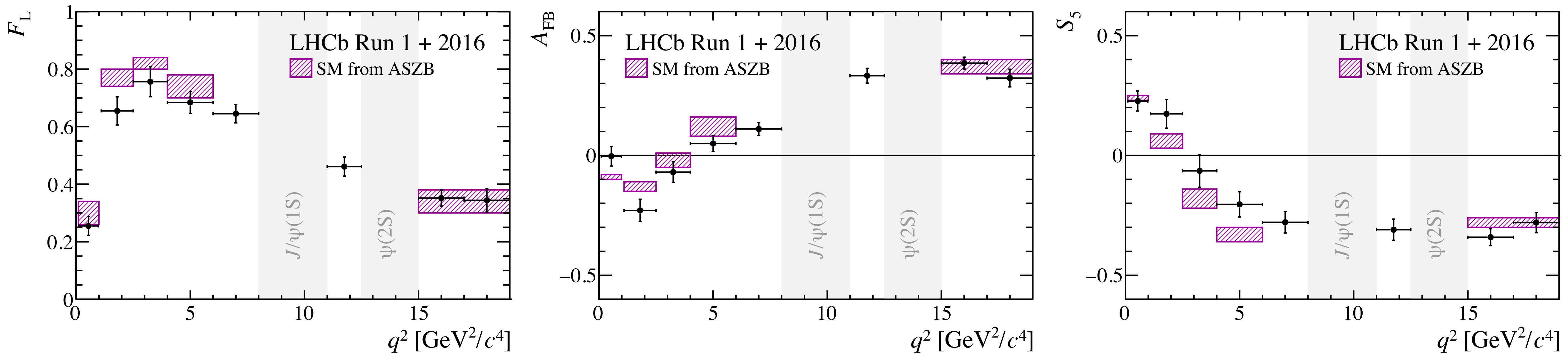


Bias correction

- Why assign a systematic uncertainty to correct for the bias seen in pseudoexperiments?
    - Bias depends on the true (unknown) value of the observables.
- Estimate the bias from the closest point in parameter space where the signal PDF is positive definite.
- The bias is typically in only one direction (to larger or to smaller values) but is assigned symmetrically in the covariance matrix.

LHCb measurements

- Full set of *CP*-averaged angular observables measured by LHCb in [PRL 125 (2020) 011802].

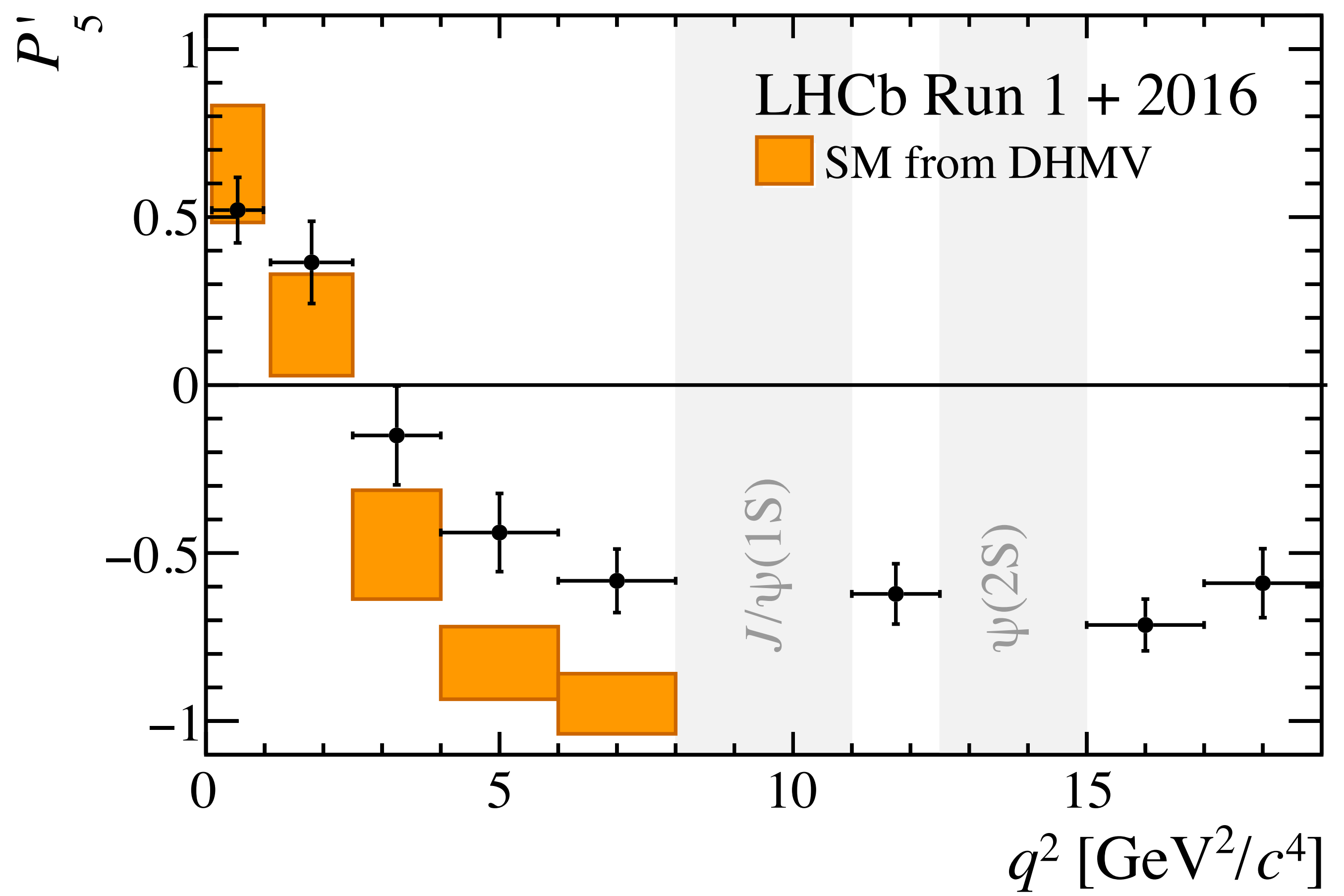


- Data compared to SM predictions based from
  [Altmannshofer & Straub, EPJC 75 (2015) 382]
  [LCSR form-factors from Bharucha, Straub & Zwicky, JHEP 08 (2016) 98]
  [Lattice form-factors from Horgan, Liu, Meinel & Wingate arXiv:1501.00367]

Why do we talk about P_5'

- In QCD factorisation/SCET there are only two independent form factors (rather than 7)
- Can construct ratios of observables which are independent of these soft form-factors at leading order, *e.g.* $P_5' = S_5/\sqrt{F_L(1-F_L)}$.

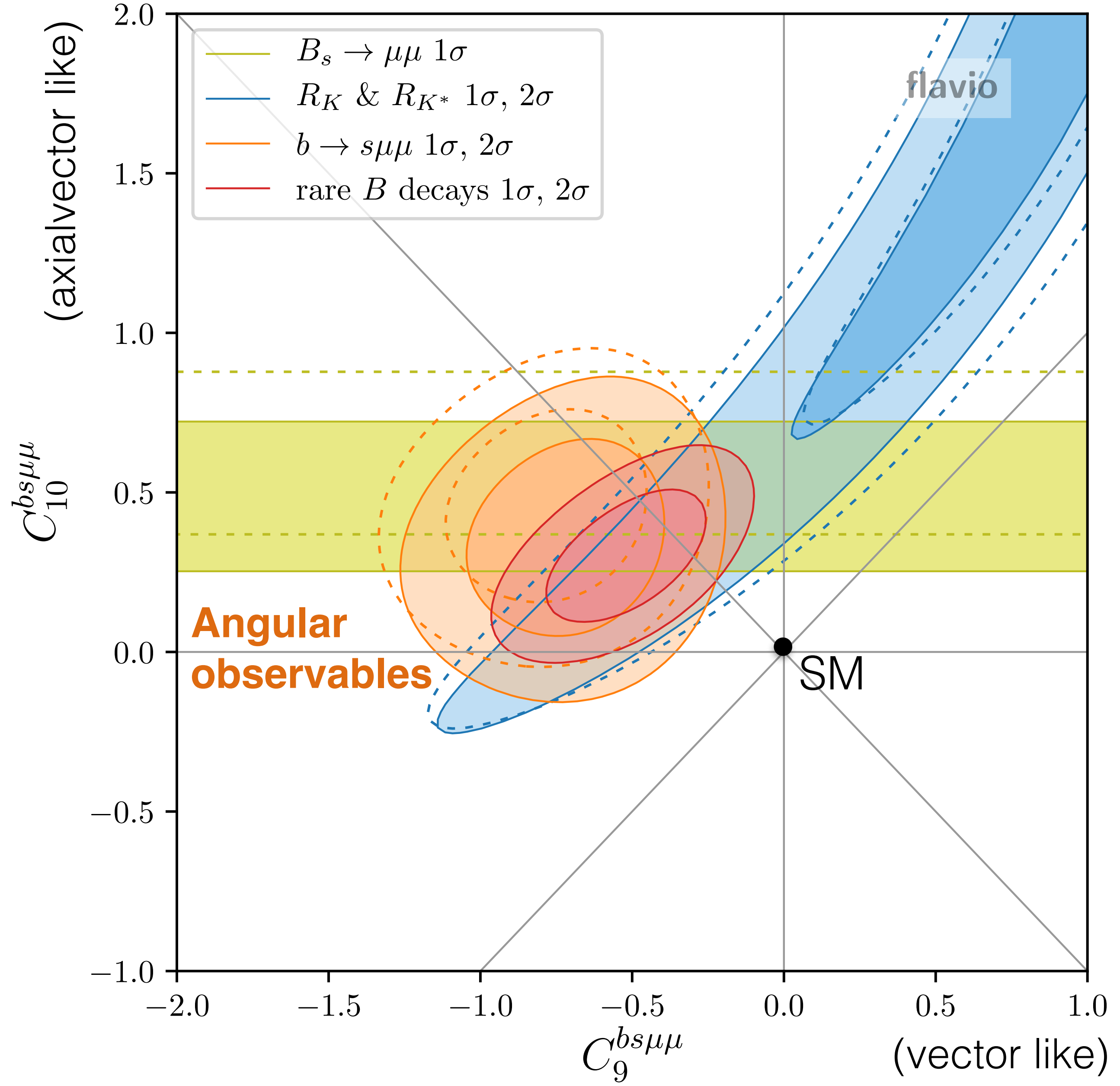
[LHCb, PRL 125 (2020) 011802]



- P_5' is one of a set of “form-factor free” observables that are measured (by reparameterising the likelihood in the fit and fitting again the data) **[Descotes-Genon et al. JHEP 1204 (2012) 104].**

Interpreting the experimental results

- Can interpret measurements of $b \to s$ processes in an effective field theory formalism.
- Same underlying Wilson coefficients appear in different decays.
- Fit all available experimental data to look for deviations from the SM.
  - Include measurements from ATLAS, BaBar, Belle, CMS and LHCb.



e.g. **[arXiv:2103.13370], [arXiv:2104.08921], [arXiv:2011.01212], [arXiv:2104.10058] …**

Global fits

- Full likelihoods are not typically available for every observable.
  - Often only available for single parameters of interest, profiling over nuisance parameters.
  - Systematic uncertainties are typically included by convoluting the experimental likelihood with a Gaussian distribution.

Challenge making the likelihood available with large number of dimensions …

Global fits

- Full likelihoods are not typically available for every observable.
  - Often only available for single parameters of interest, profiling over nuisance parameters.
  - Systematic uncertainties are typically included by convoluting the experimental likelihood with a Gaussian distribution.
- Instead, treat experimental measurements using a multivariate Gaussian likelihood:

Parameters of interest

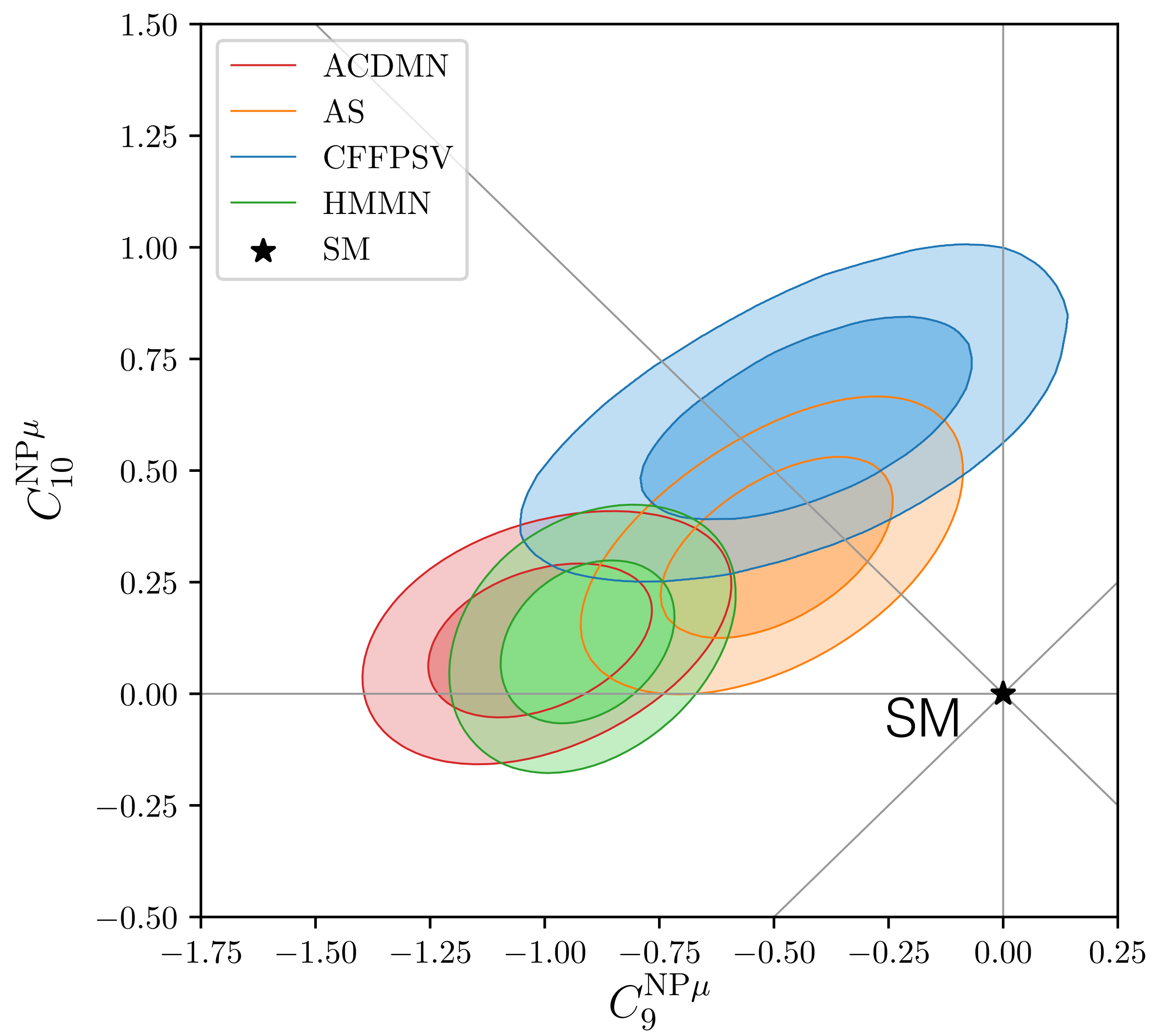
$$L_{\mathrm{exp}} = \exp\left(-\frac{1}{2}(\vec{a}(\theta_{\mathrm{th}}) - \vec{a})^{\mathrm{T}} C^{-1} (\vec{a}(\theta_{\mathrm{th}}) - \vec{a})\right)$$

- Constructing a covariance matrix $C = C_{\mathrm{stat}} + C_{\mathrm{syst}}$

Assumes systematic uncertainties are normally distributed

Global fits

- Different fitting groups use different experimental inputs, SM assumptions and statistical treatments ($\Delta\chi^2$ vs bayesian analysis).



[AS, arXiv:2103.13370]
[ACDMN, arXiv:2104.08921]
[CFFPSV, arXiv:2011.01212]
[HMMN, arXiv:2104.10058]

Discrepancies are due to the choice of experimental inputs and treatment of systematic uncertainties on theoretical predictions.

See talk by P. Stangl et al. at the "Flavour Anomaly Workshop 2021"

Summary

- Flavour physics encompasses a very broad range of different measurements, each with different sources systematic uncertainty.
- Typically consider systematic uncertainties:
  - From statistical uncertainties on external inputs.
  - Due to biasing effects in our measurement.
- We are often only able to consider a single variation out of many possible variations due to practical constraints.

Likelihood

- Likelihood for a single data set:

B^0 mass distribution, distinguishes signal from background

Signal angular distribution

Efficiency

$$L = \left[\prod_i^N \frac{1}{N_S+N_B}\left(N_S\epsilon_S(\overrightarrow{\Omega}^i)P_S(\overrightarrow{\Omega}^i)P_S(m^i_{K\pi\mu\mu}) + N_B P_B(\overrightarrow{\Omega}^i)P_B(m^i_{K\pi\mu\mu})\right)\right] \times$$

$$\left[\prod_i^N \frac{1}{N_S+N_B}\left(N_S P_S(m^i_{K\pi}) + N_B P_B(m^i_{K\pi})\right)\right] \times$$

$$P(N|N_S+N_B)$$

Background

$K^+\pi^-$ mass to separate S- and P-wave K^* states

Extended (Poisson) term

Operators

- Different processes are sensitive to different 4-fermion operators.
  - ➡ Can exploit this to over-constrain the system.

$$\mathcal{O}_7 = (m_b/e)\,(\bar{s}\sigma^{\mu\nu}P_R b F_{\mu\nu})$$
$$\mathcal{O}_9 = (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \ell)$$
$$\mathcal{O}_{10} = (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \gamma_5 \ell)$$
$$\mathcal{O}_S = (\bar{s}P_R b)(\bar{\ell}\ell)$$
$$\mathcal{O}_P = (\bar{s}P_R b)(\bar{\ell}\gamma_5 \ell)$$

photon (constrained by radiative decays and $b \to s\ell^+\ell^-$ processes at small q^2)

vector current (constrained by $b \to s\ell^+\ell^-$ processes)

axial vector current (constrained by leptonic decays and $b \to s\ell^+\ell^-$ processes)

scalar and pseudoscalar operators (constrained primarily by leptonic decays)

e.g. $B_s^0 \to \mu^+\mu^-$ constrains $C_{10} - C'_{10},\ C_S - C'_S,\ C_P - C'_P$

$B^+ \to K^+\mu^+\mu^-$ constrains $C_9 + C'_9,\ C_{10} + C'_{10}$

$B^0 \to K^{*0}\mu^+\mu^-$ constrains $C_7 \pm C'_7,\ C_9 \pm C'_9,\ C_{10} \pm C'_{10}$

The primes denote right-handed counterparts of the operators whose contribution is small in the SM.

Effective theory

- Can write a Hamiltonian for an effective theory of $b \to s$ processes:

Wilson coefficient
(integrating out scales above μ)

Local 4 fermion operators with different Lorentz structures

$$\mathcal{H}_{\mathrm{eff}} = -\frac{4\,G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{\alpha_e}{4\pi} \sum_i \boxed{C_i(\mu)} \boxed{\mathcal{O}_i(\mu)},$$

$$\Delta\mathcal{H}_{\mathrm{eff}} = \frac{\kappa_{\mathrm{NP}}}{\Lambda^2_{\mathrm{NP}}} \boxed{\mathcal{O}_{\mathrm{NP}}}$$

NP scale

NP can modify SM contribution or introduce new operators

- κ_{NP} can have all/some/none of the suppression of the SM, *e.g.* MFV inherits SM CKM suppression.

c.f. Fermi theory of weak interaction where at low energies:

$$\lim_{q^2 \to 0} \left(\frac{g^2}{m_W^2 - q^2} \right) = \frac{g^2}{m_W^2}$$

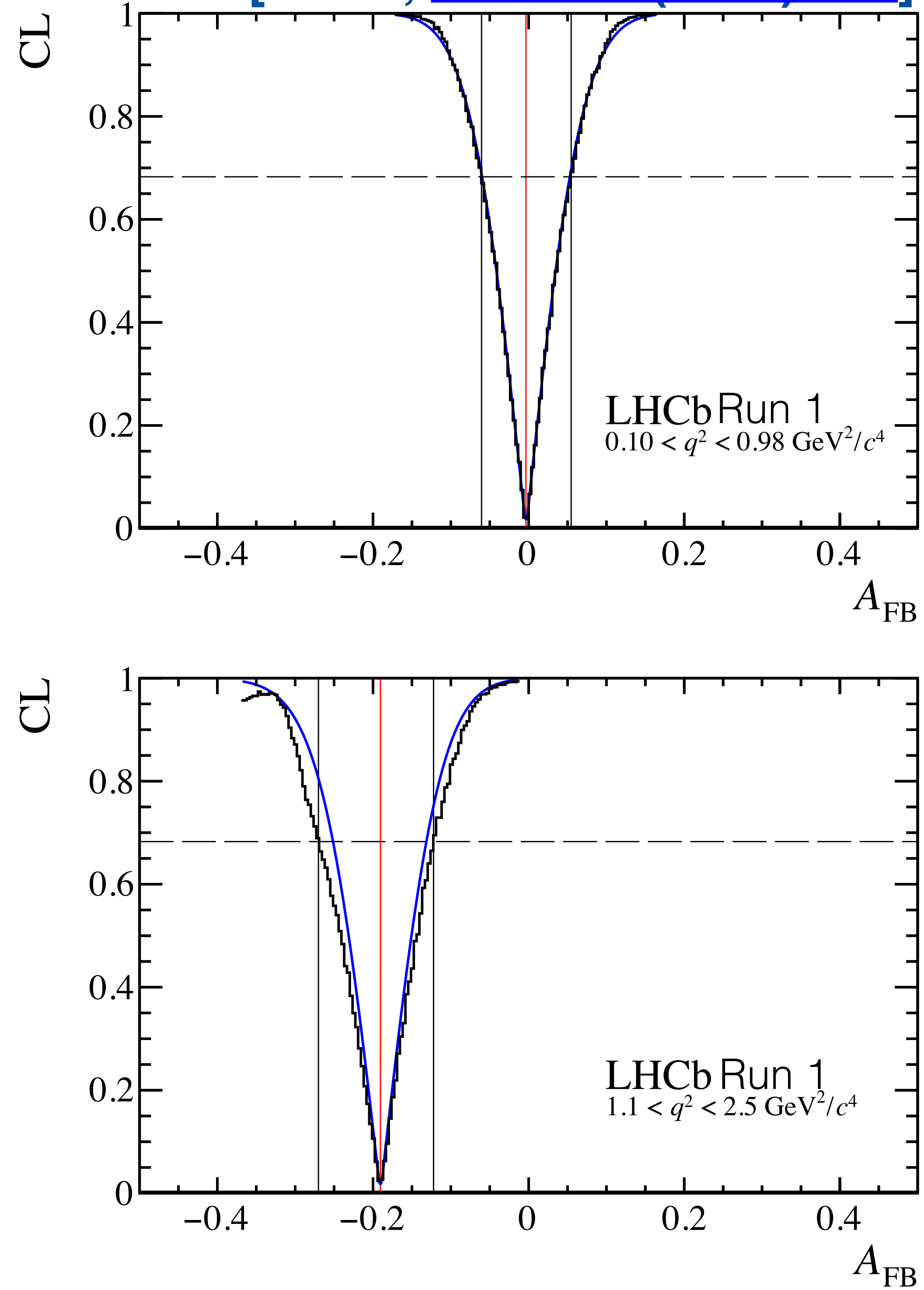
i.e. the full theory can be replaced by a 4-fermion operator and a coupling constant, G_{F}

Statistical uncertainties in Run 1

- Estimate of covariance matrix from the Hessian appears to have correct coverage when fitting the combined Run 1 + 2016 data sets.
- Problems are seen when fitting to single data taking period.
- Previously used the Feldman Cousins **[PRD 57 (1998) 3873-3889]** approach to construct intervals for the Run 1 data set.
    - Neyman construction with the likelihood ratio as an ordering principle.
- Use plug-in method for treatment of nuisance parameters in pseudoexperiment generation (from the best-fit point in data).
  [B. Sen, M. Walker & M. Woodroofe. Stat. Sinica 19 301].

Statistical uncertainties in Run 1

[LHCb, JHEP 02 (2016) 104]



- Compare confidence intervals from the profile likelihood and Feldman Cousins interval.
- Differences seen in q^2 bins with lowest signal yields.

Statistical uncertainties in Run 1

[LHCb, JHEP 02 (2016) 104]

- Largest effects are seen in the P_i observables in the bin where F_{L} is largest (and the yield is small).