INVESTIGATING THE FLAVOR ANOMALIES

MARTA ANDRÉ, REZA JAFARI, RUBEN POZZI

LIP INTERNSHIP PROGRAM 2021

SUPERVISORS: ALESSIO BOLETTI, MARIA FARIA, NUNO LEONARDO

OUTLINE OF THE PRESENTATION

- **Introduction & goal**
- **Steps of the project:**
  1. **Fit the inv. mass of B0 with systematic & statistical uncertainty estimation**
  2. **Estimate efficiency**
  3. **Measure branching fraction**
- **Analysis of the b-parking dataset (ongoing)**

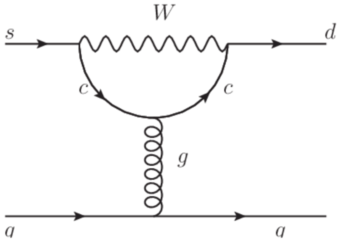
INTRODUCTION
FLAVOR ANOMALIES

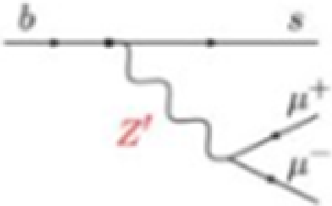
- **A possible violation of flavor universality.**
- **Deviation from the SM prediction in the flavor sector .**
- **One of the most interesting areas at the LHC.**
- **Most significant deviation from SM.**

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

b→sll decays

- **Focus on: Flavor Rare Decays**
  - **FCNC : penguin diagrams, …**

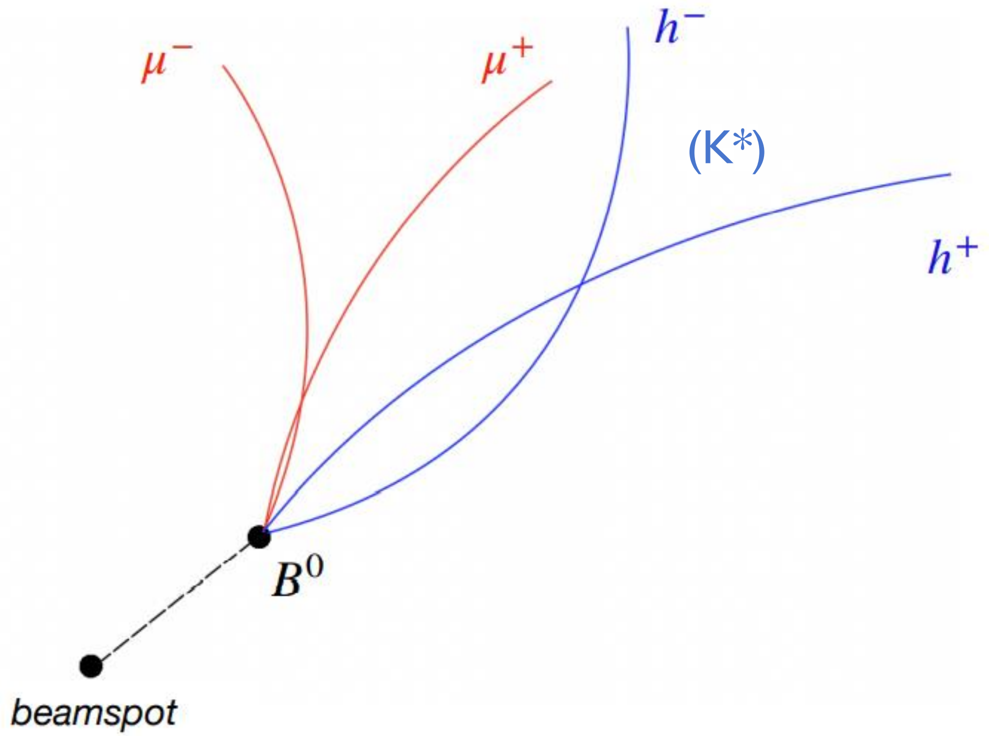




INTRODUCTION
THE GOAL

- Investigate the flavor anomalies by analyzing the run-2 pp data from the years 2016, 2017, 2018, total integrated luminosity of 139.5 fb^{-1} at $\sqrt{s}$ = 13 TeV.
- Fit the invariant mass spectrum of $\mathrm{B}^0 \to K^{*0}\ \mu^+\mu^-$, estimate the systematic uncertainties and the efficiency.
- **Use all of these to measure the branching fraction of the decay:**

$$\frac{d\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{dq^2} = \frac{Y_S}{Y_N}\frac{\epsilon_N}{\epsilon_S}\frac{\mathcal{B}(B^0 \to K^{*0}J/\psi)}{\Delta q_i^2}$$



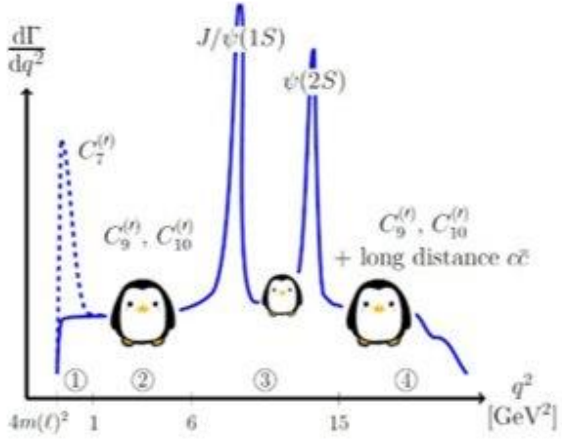
INTRODUCTION
THE ANALYSIS

- **The data is divided in bins of the di-muon invariant mass squared (q2)**

- **Independent analysis of each bin.**

- **B0 → J/ψ K*0 (bin4) is used as normalization channel**

- **Right-tagged (RT) events and wrong-tagged events (WT) are included because 12-14% of the events have a wrong Kπ flavor assignment for the K*0 invariant mass**



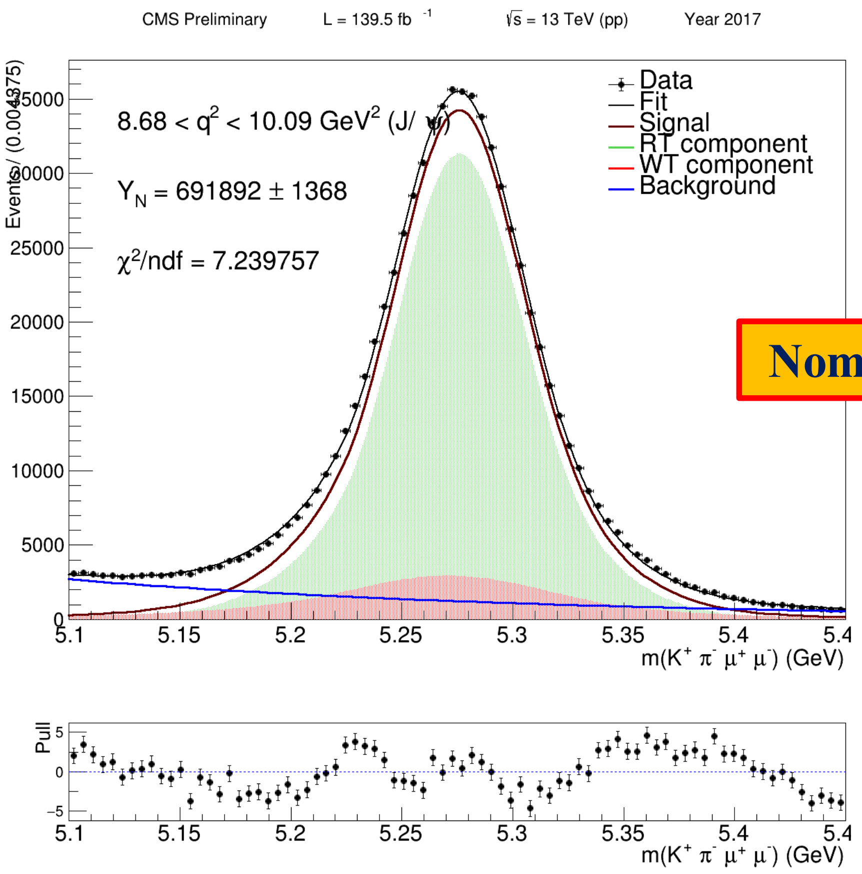
Bin index	q^2 range [GeV²]
0	1-2
1	2-4.3
2	4.3-6
3	6-8.68
4 J/ψ	8.68-10.09
5	10.09-12.86
6 $\psi(2S)$	12.86-14.18
7	14.18-16

HTTPS://ESPACE.CERN.CH/CMS-LIP/INTERNSHIP/SUMMER2021/ANOMALIES/INTRO%20MATERIAL/B0TOKSTARMUMU-ANALYSIS-INTRO-FALL2020.PDF

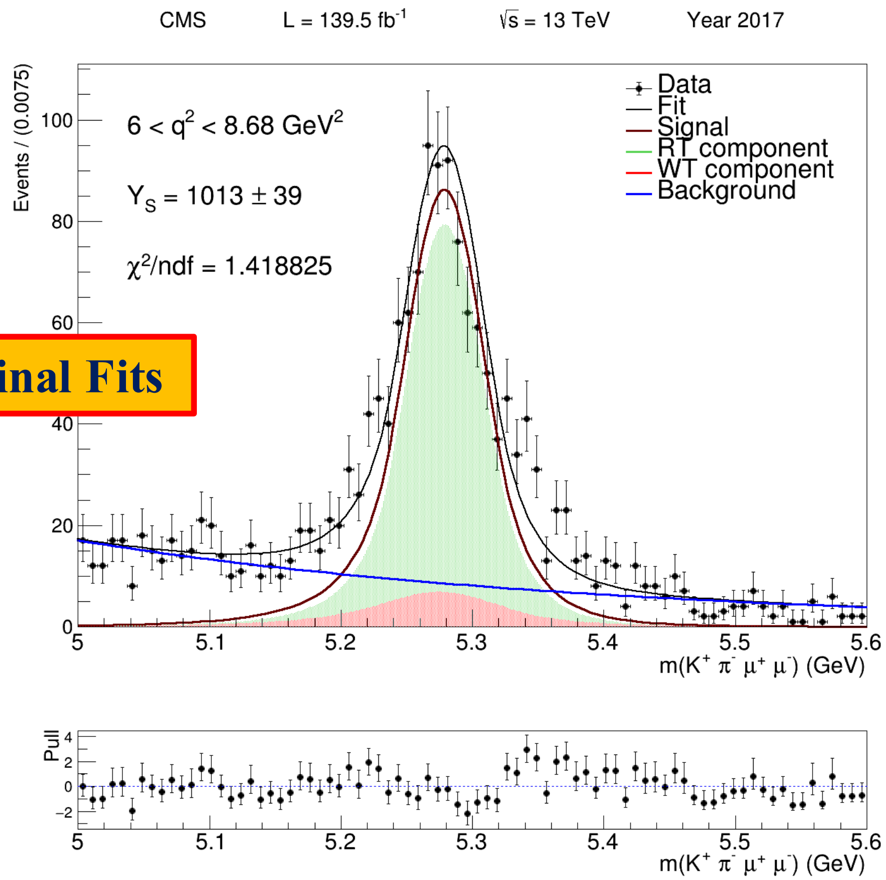
STEP ONE: FIT THE MASS OF B0

THE FIT OF SIGNAL BIN (2017)

- RooFit is the main tool used in the program, which is used to fit the datasets to custom functions and parameters
- RT MC and WT MC are fitted separately and then used in the fits on data.
- Data is fitted applying Gaussian constraints on all fit parameters using the values obtained from MC



Bin = 4 (normalization channel)



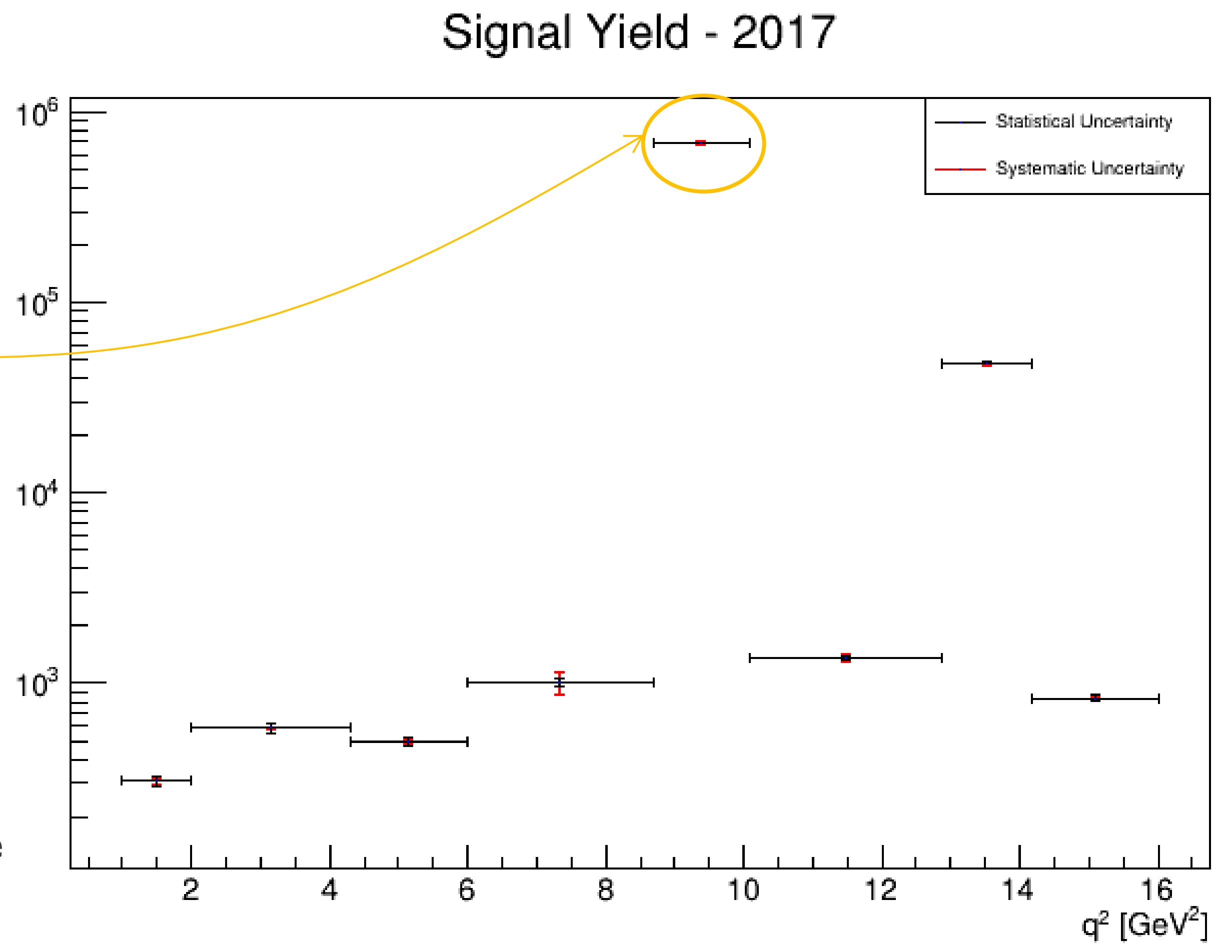
Nominal Fits

Bin = 3 (signal channel)

STEP ONE: FITTING THE MASS OF B0

EXTRACT YIELD AND UNCERTAINTIES

$$\frac{d\mathscr{B}(B^0 \to K^{*0}\mu^+\mu^-)}{dq^2} = \frac{Y_S}{Y_N}\frac{\epsilon_N}{\epsilon_S}\frac{\mathscr{B}(B^0 \to K^{*0}J/\psi)}{\Delta q_i^2}$$



- **Yield (number of events) is extracted from the fit.**
- **Removing constraints on the data and compare to nominal fit to find the fit variation**
- **Systematics are calculated by** $syst = \frac{abs(N_1 - N_0)}{N_0}$, **where** N_1 **is the yield of the fit variation and** N_0 **is the nominal yield.**

STEP TWO: EFFICIENCY

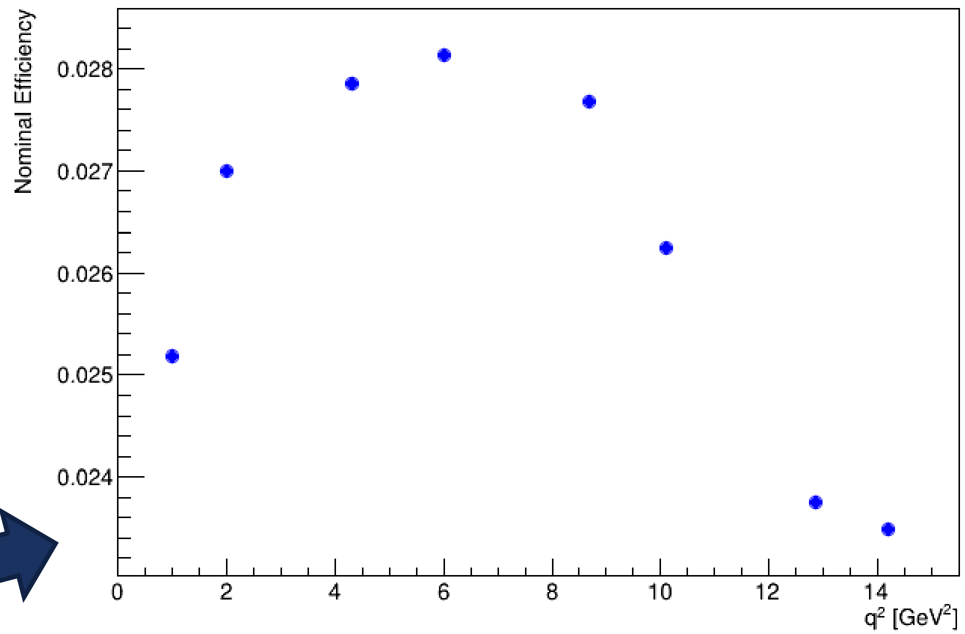
BRANCHING FRACTION EFFICIENCY

$$\frac{d\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{dq^2} = \frac{Y_S}{Y_N}\frac{\epsilon_N}{\epsilon_S}\frac{\mathcal{B}(B^0 \to K^{*0}J/\psi)}{\Delta q_i^2}$$

ε_N – normalization efficiency

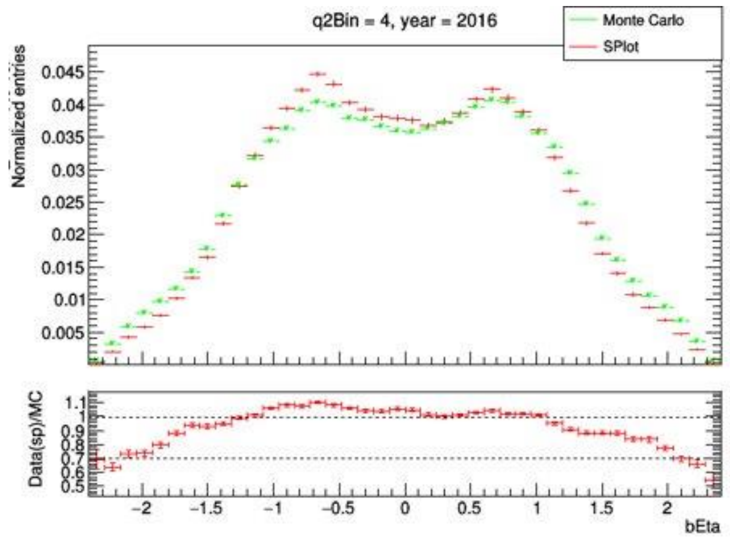
ε_S – signal efficiency

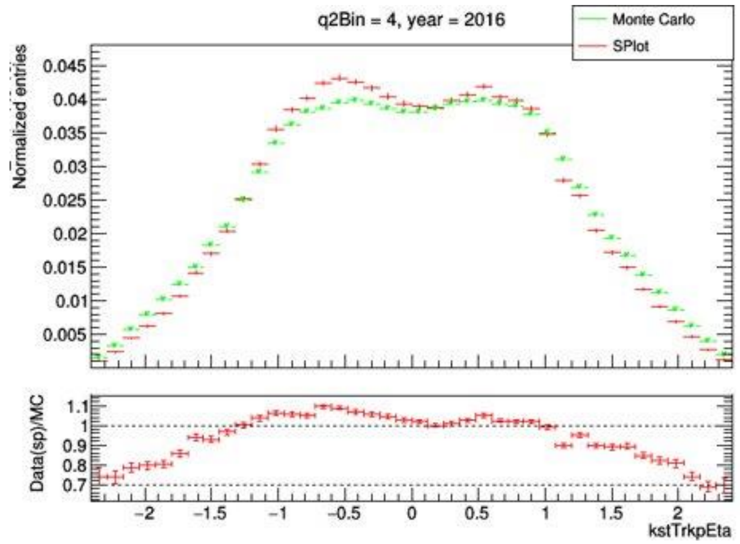
- Monte Carlo fit gives output of experiment not influenced by detector;
- Compute the efficiencies with Monte Carlo fit.

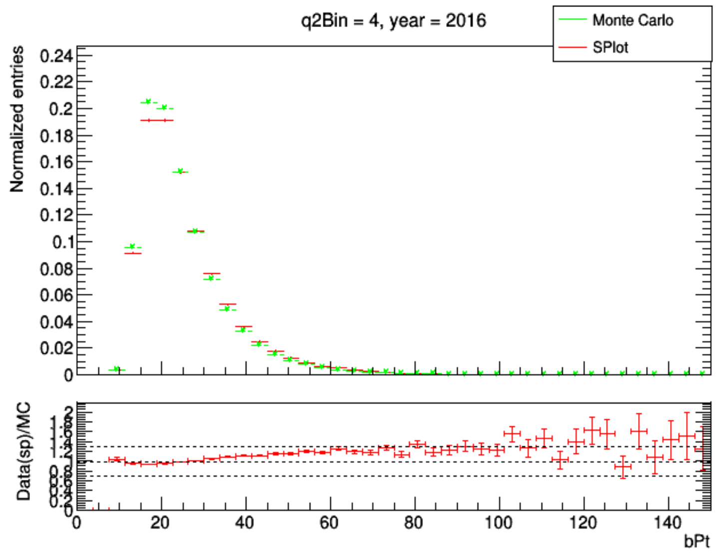


STEP TWO: EFFICIENCY
WEIGHTING EFFICIENCY

- Separate signal from background for data fit;
- Compare signal events between the MC fit and the data fit;
- The data-MC ratio is used to re-weight the MC-derived efficiencies.



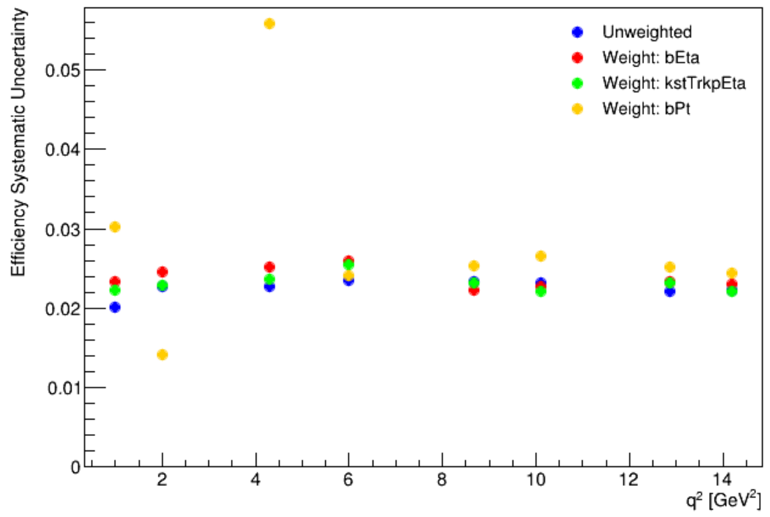




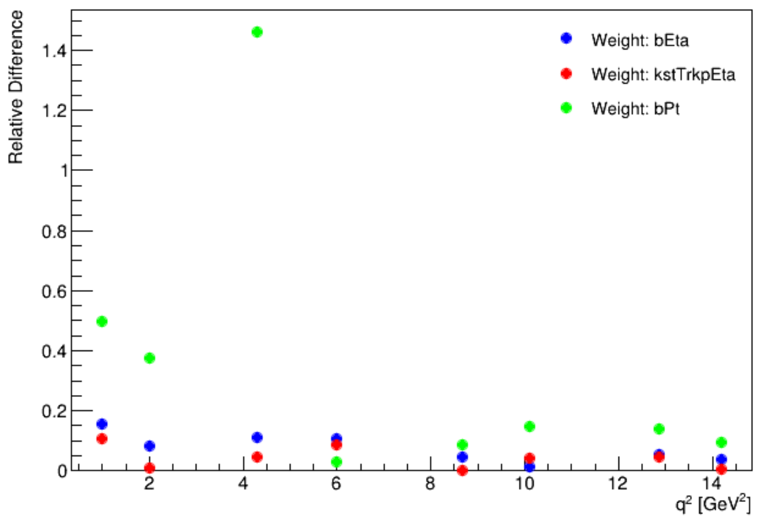
STEP TWO: EFFICIENCY

EFFICIENCY VS. WEIGHTED EFFICIENCY AND RELATIVE DIFFERENCE

- **Efficiency:** Efficiency computed using Monte Carlo events; → $\epsilon = \frac{N_{selected}}{N_{all}}$
- **Weighted efficiency:** Correction made to the efficiency through the data/MC ratio. → $\epsilon^{wei} = \frac{N_{acc}}{D_{acc}} \times \frac{N_{sel}^{wei}}{D_{sel}^{wei}}$



Year: 2016



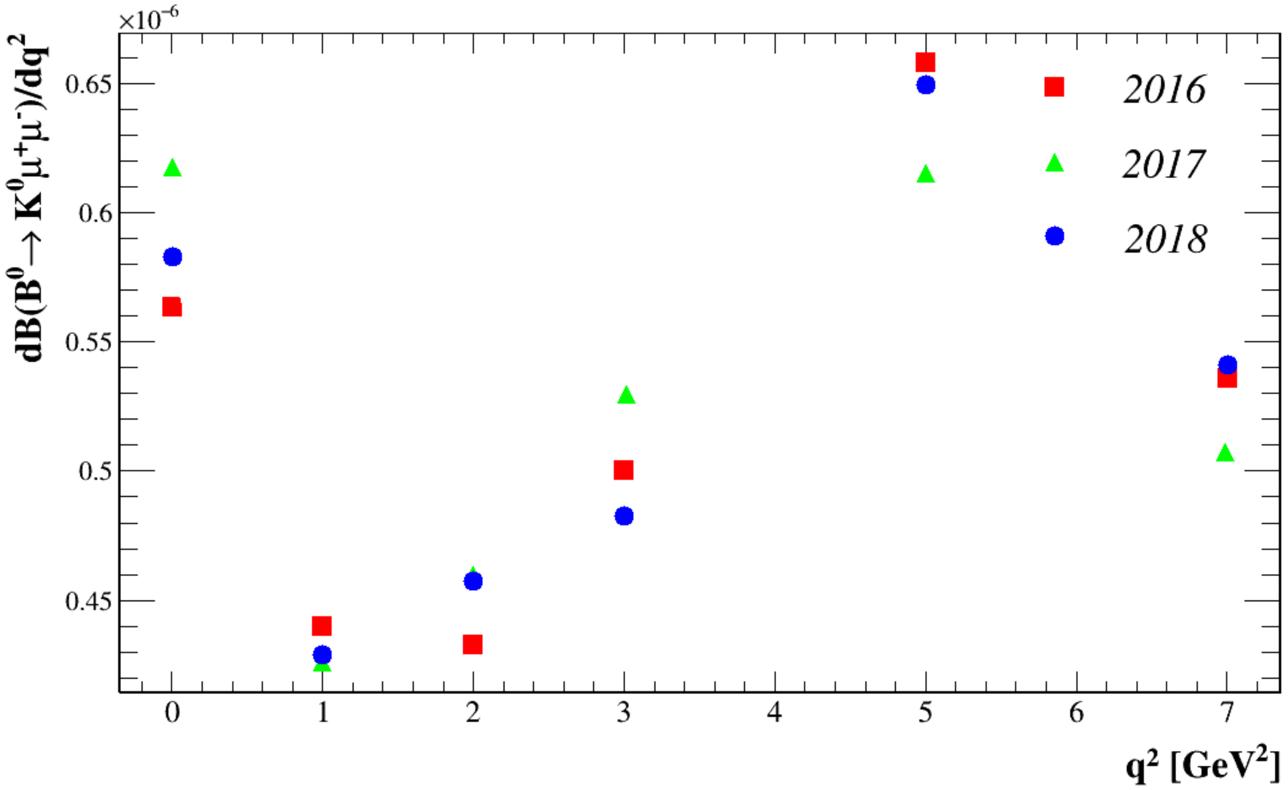
Year: 2016

STEP THREE: BRANCHING FRACTION
MEASUREMENT OF BF

BF for the Signal channels: Now, we have the parameters below:

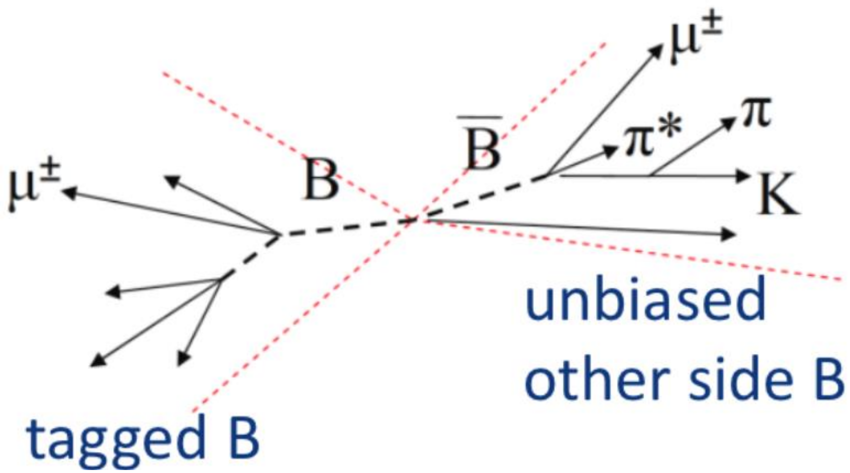
$$\frac{d\mathscr{B}(B^0 \to K^{*0}\mu^+\mu^-)}{dq^2} = \frac{Y_S}{Y_N}\frac{\epsilon_N}{\epsilon_S}\frac{\mathscr{B}(B^0 \to K^{*0}J/\psi)}{\Delta q_i^2}$$

- Yields (N , S);
- Efficiencies (N, S);
- BF for the normalization channel;
- Width of q2 bins.



ONGOING PROJECT: B-PARKING ANALYSIS
B-PARKING DATASET

- Dataset of 12 billion events that is enriched (80%) by events where a pair of B-hadrons are produced.
  - One of the B in the pair has trigger requirement that one muon is produced
  - The other B has no trigger selection (unbiased), which means they can decay into anything
- B-parking is used when we can't develop a trigger for a specific decay (i.e., low pT events with no muons)
- Special data stream collected by CMS to explore flavor anomalies

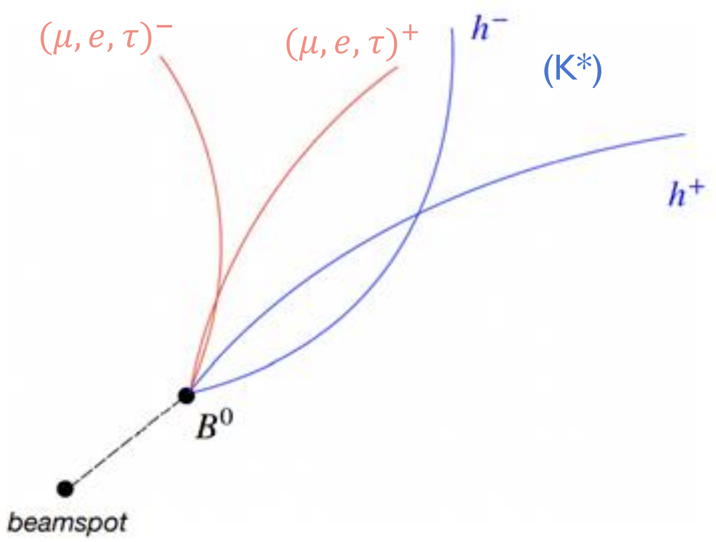


HTTPS://INDICO.CERN.CH/EVENT/744951/CONTRIBUTIONS/3171458/ATTACHMENTS/1741378/2817510/LLP_SWAGATA.PDF

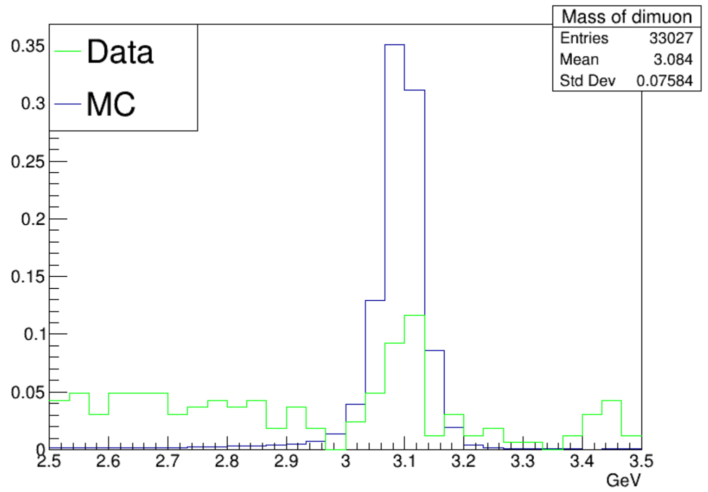
B-PARKING ANALYSIS
THE ANALYSIS

1. Reconstruct the candidates for the $\mathrm{B}^0 \to K^{*0}\ \mu^+\mu^- \to K^+\pi^-\mu^+\mu^-$
2. Develop machine learning selections to reject the background
3. Select the fitting function by fitting MC samples
4. Fit data samples to extract the signal yield and compare it with the yield of the non-B-parking analysis

- This analysis on the b-parking data are not expected to be more precise than the analysis on normal data
  - this decay can be used as normalization channel for the analysis of the same decay with other leptons
  - e.g. $(B^0 \to K^{*0}ee$ or $B^0 \to K^{*0}\tau\tau)$
  - Helpful to study lepton-flavor universality

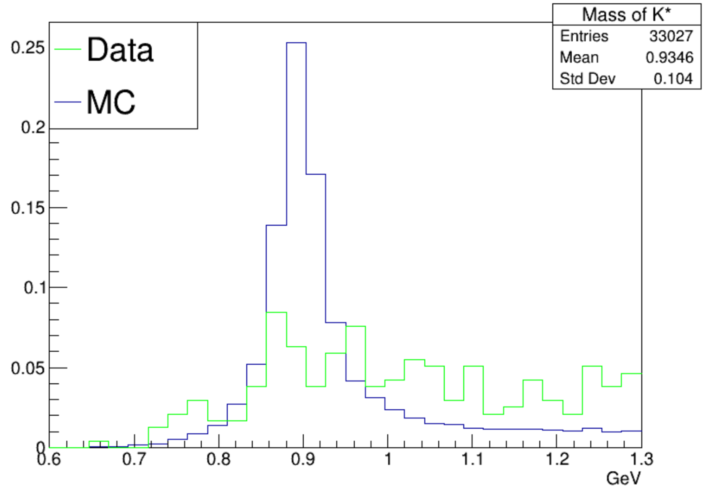


B-PARKING ANALYSIS
SELECTION CRITERIA



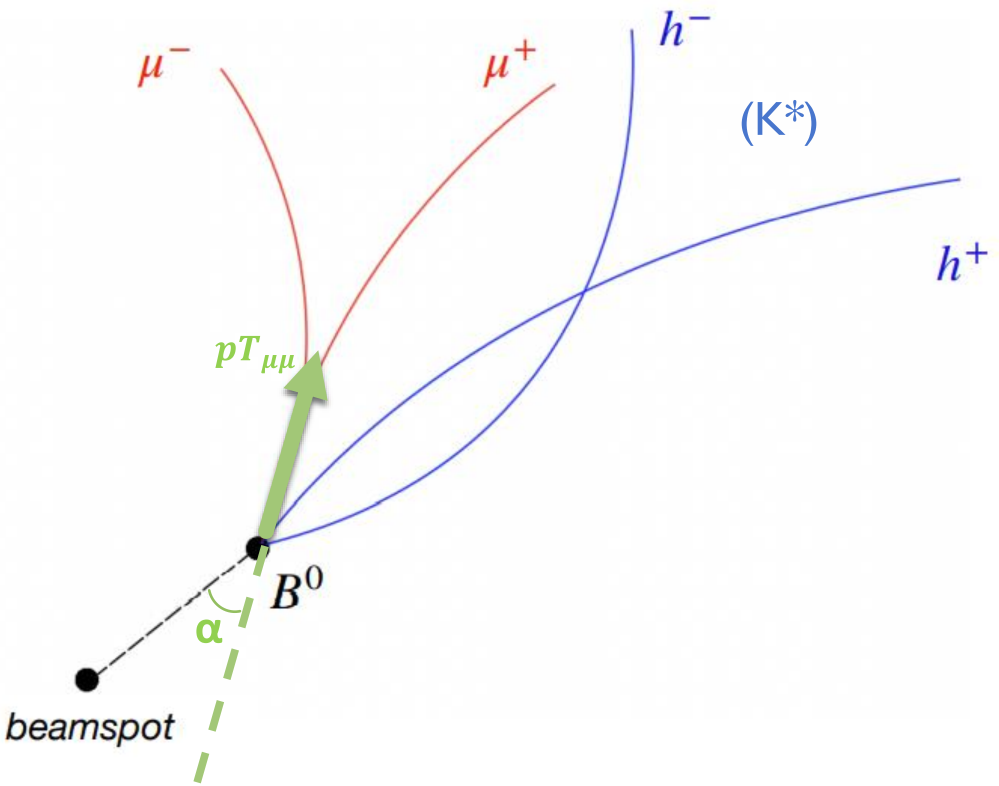
Dimuon selection cuts:

- Mass: 3 – 3.15 GeV
- $\cos(\alpha)$: >0.9



K selection cuts:*

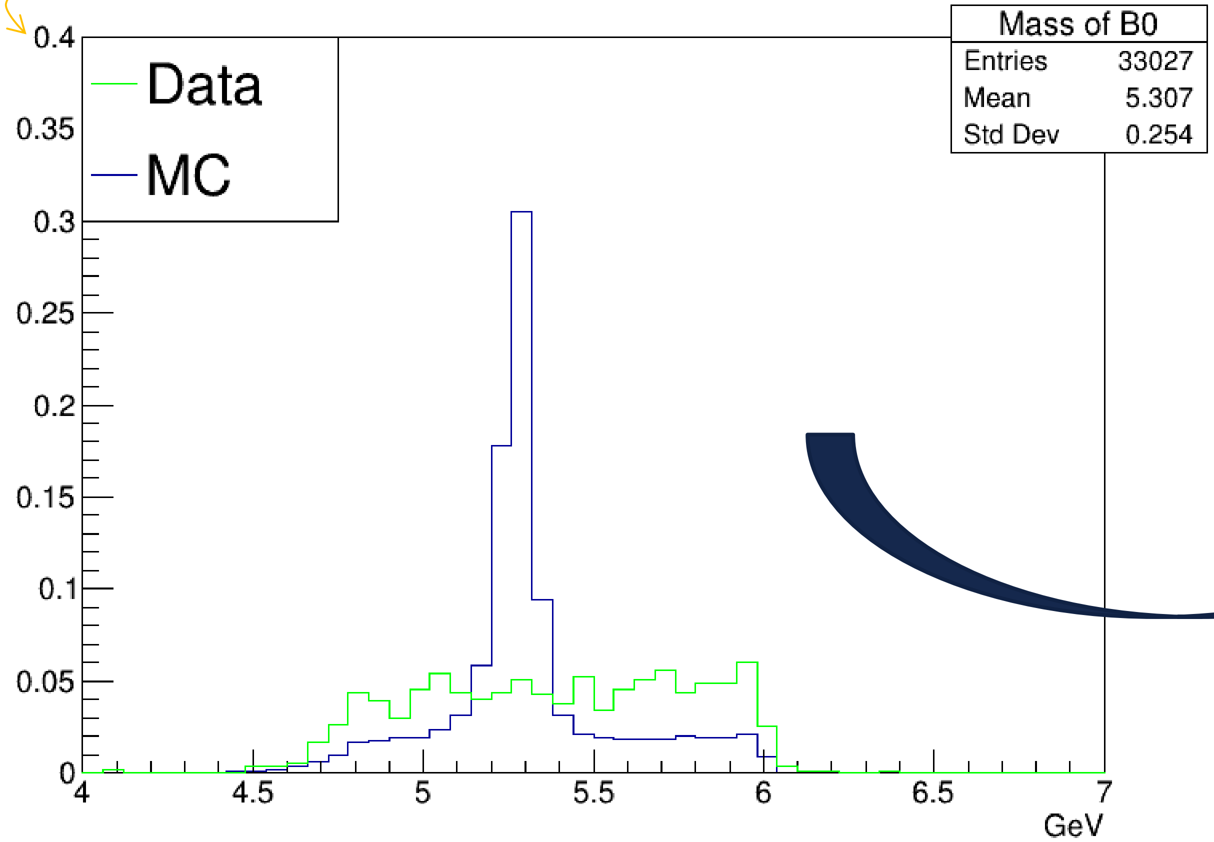
- Mass: 0.8 – 1 GeV



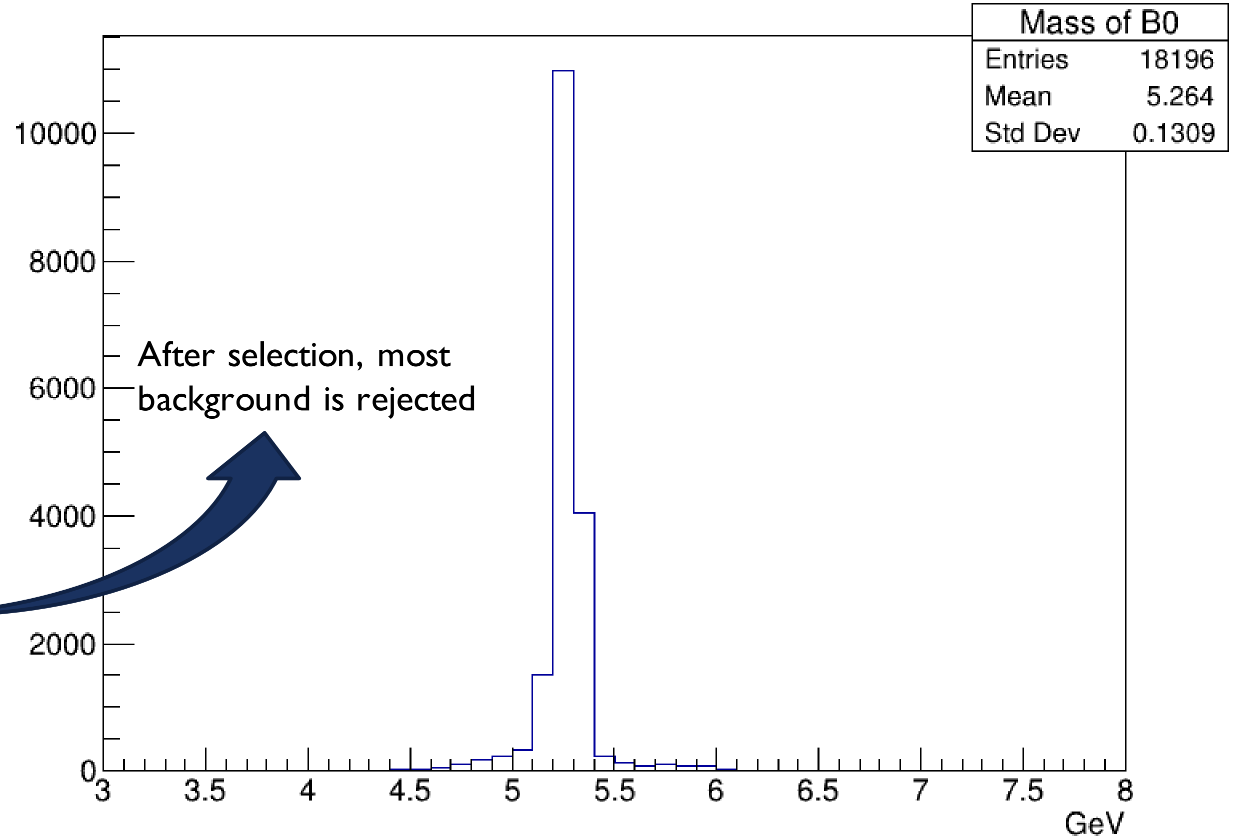
B-PARKING ANALYSIS

B^0 MASS AFTER SELECTION

Normalized to show non-MC data



After selection, most background is rejected



CONCLUSION
SUMMARY & FUTURE OUTLOOK

- In this project we have gone through three main steps:
  1. Fitted the invariant mass squared of the B0 to study its decay
  2. Estimated the systematic uncertainties and the efficiencies
  3. Measured the branching fraction of the decay
- Fascinating overview of the flavor anomalies
- Learnt a lot of the theoretical background and how to use new tools such as RooFit
- Insight on how analysis of the huge amount of data from LHC is carried out

What's next?

- The B-parking project ongoing until the half of October
- Extend the analysis to a new dataset and compare b-parking results with previous analysis

THANK YOU FOR LISTENING!

QUESTIONS?

BACKUP

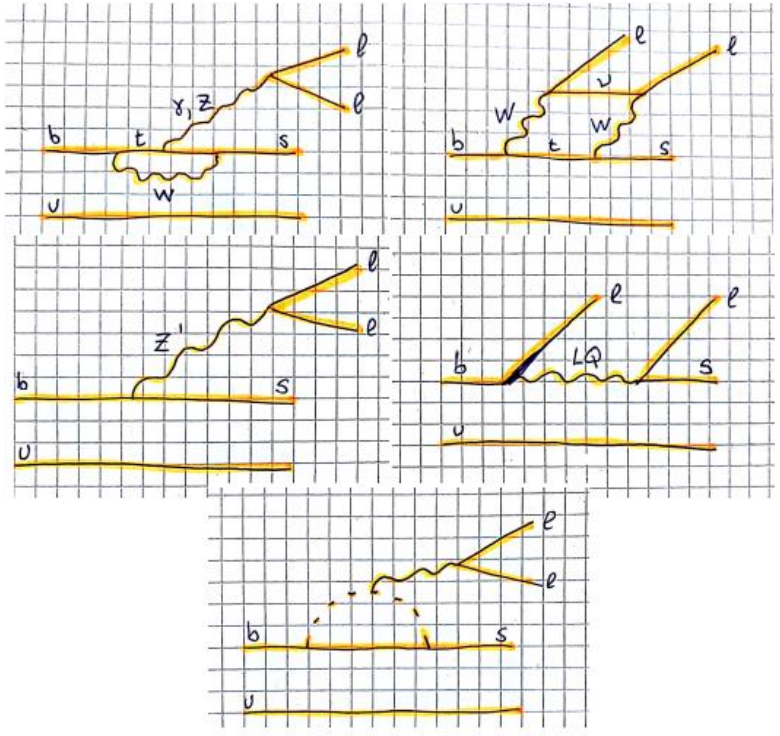
SYSTEMATICS

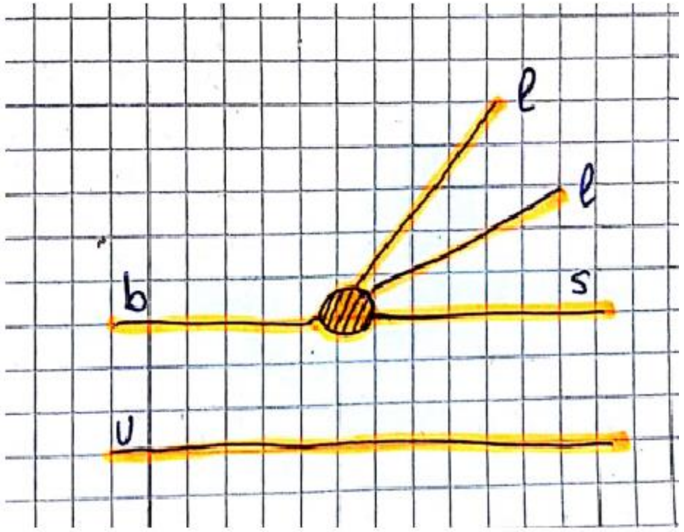
Table 1: Systematics for the fit variation, where $syst = \frac{abs(N_1 - N_0)}{N_0}$, and N_1 is the yield for the new variation and N_0 the nominal one

q^2 bin	Year	Nominal yield	Variation yield	Systematics (relative)
0	2016	199 ± 17 (0.085)	192 ± 19 (0.097)	0.033
	2017	307 ± 21 (0.069)	303 ± 23 (0.076)	0.014
	2018	501 ± 27 (0.054)	468 ± 29 (0.061)	0.064
1	2016	402 ± 26 (0.064)	391 ± 27 (0.069)	0.026
	2017	581 ± 30 (0.051)	580 ± 32 (0.055)	0.003
	2018	982 ± 39 (0.040)	938 ± 42 (0.045)	0.045
2	2016	341 ± 24 (0.072)	339 ± 27 (0.079)	0.005
	2017	495 ± 28 (0.056)	480 ± 29 (0.061)	0.030
	2018	822 ± 36 (0.044)	779 ± 39 (0.050)	0.052
3	2016	667 ± 33 (0.049)	630 ± 36 (0.057)	0.055
	2017	1013 ± 39 (0.039)	940 ± 42 (0.044)	0.072
	2018	1609 ± 50 (0.031)	1503 ± 53 (0.035)	0.066
4	2016	492337 ± 1206 (0.002)	492389 ± 1212 (0.002)	0.000
	2017	691892 ± 1368 (0.002)	691818 ± 1370 (0.002)	0.000
	2018	1427613 ± 2109 (0.001)	1422529 ± 85 (0.000)	0.004
5	2016	1087 ± 42 (0.039)	1148 ± 50 (0.044)	0.055
	2017	1363 ± 46 (0.034)	1430 ± 53 (0.037)	0.049
	2018	3022 ± 68 (0.022)	3157 ± 77 (0.024)	0.045
6	2016	37533 ± 268 (0.007)	36077 ± 268 (0.007)	0.039
	2017	48127 ± 324 (0.007)	48630 ± 1080 (0.022)	0.010
	2018	98085 ± 451 (0.005)	93873 ± 436 (0.005)	0.043
7	2016	804 ± 63 (0.078)	676 ± 31 (0.045)	0.159
	2017	836 ± 33 (0.040)	833 ± 35 (0.043)	0.003
	2018	1837 ± 49 (0.027)	1805 ± 52 (0.029)	0.018

- **Effective Field Theory (EFT):**
  - **A powerful tools in searching for new physics**
  - **Model-independent approach**

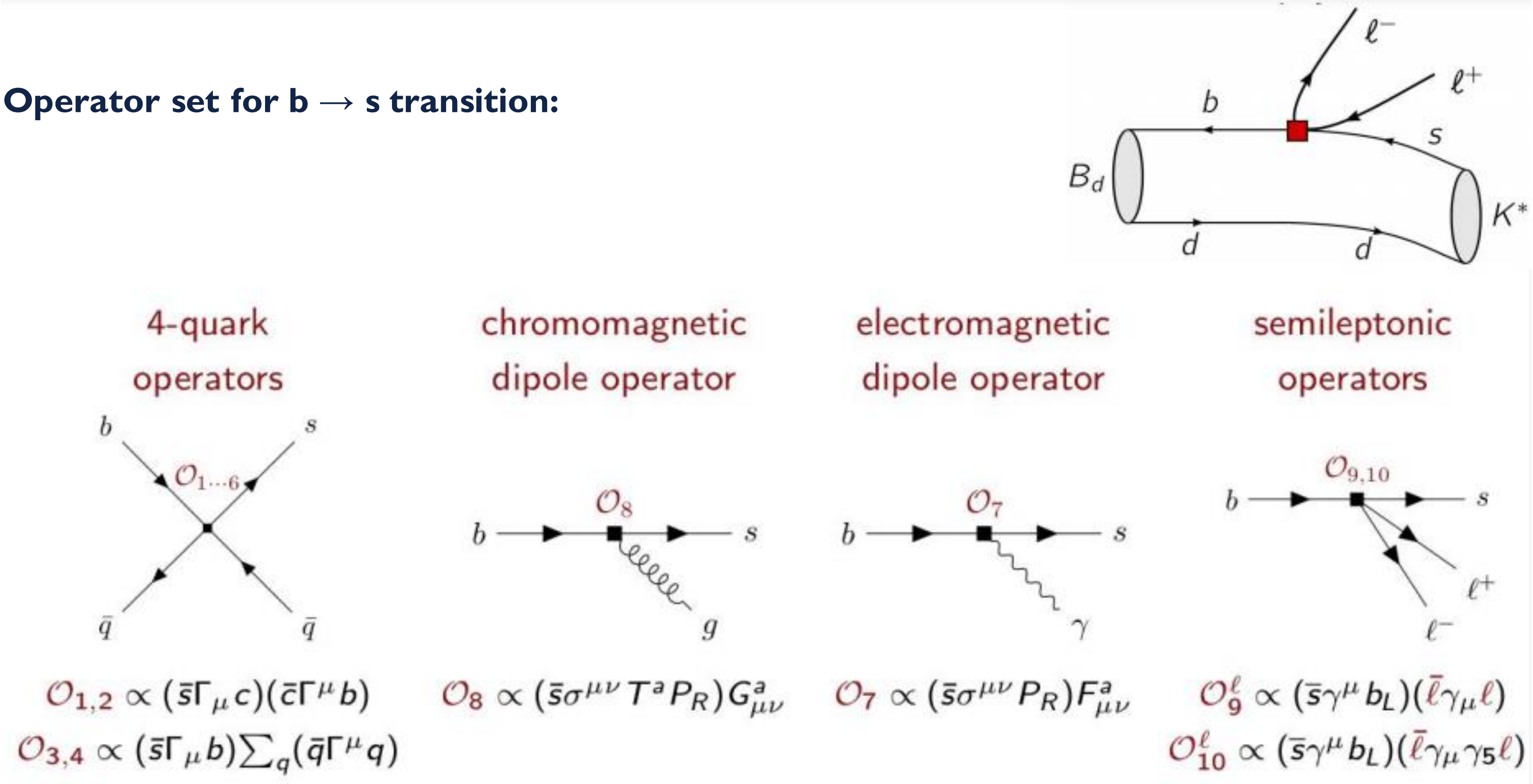
$$\mathcal{L} \propto \sum_i \mathcal{C}_i \mathcal{O}_i, \qquad \mathcal{C}_i \equiv \mathcal{C}_i^{SM} + \mathcal{C}_i^{NP} \qquad \text{"Wilson coefficients"}$$





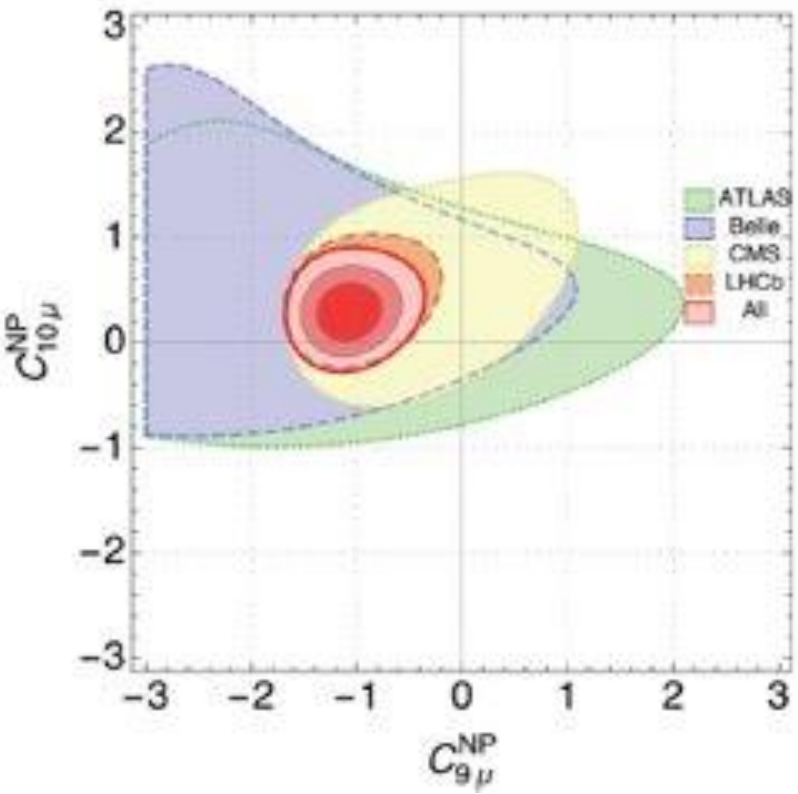
EFT Operators

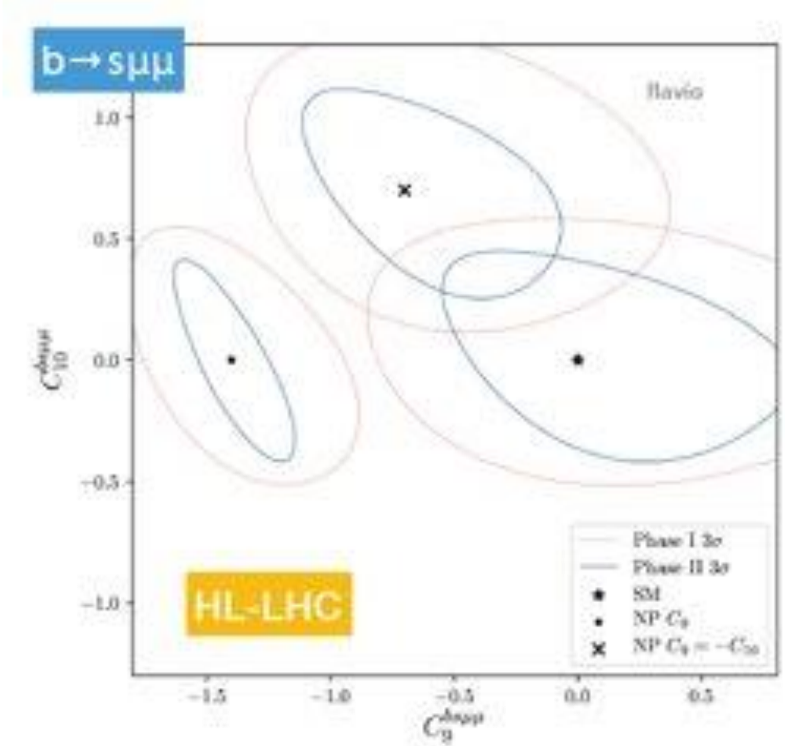
- **Operator set for b → s transition:**



SOME RESULTS

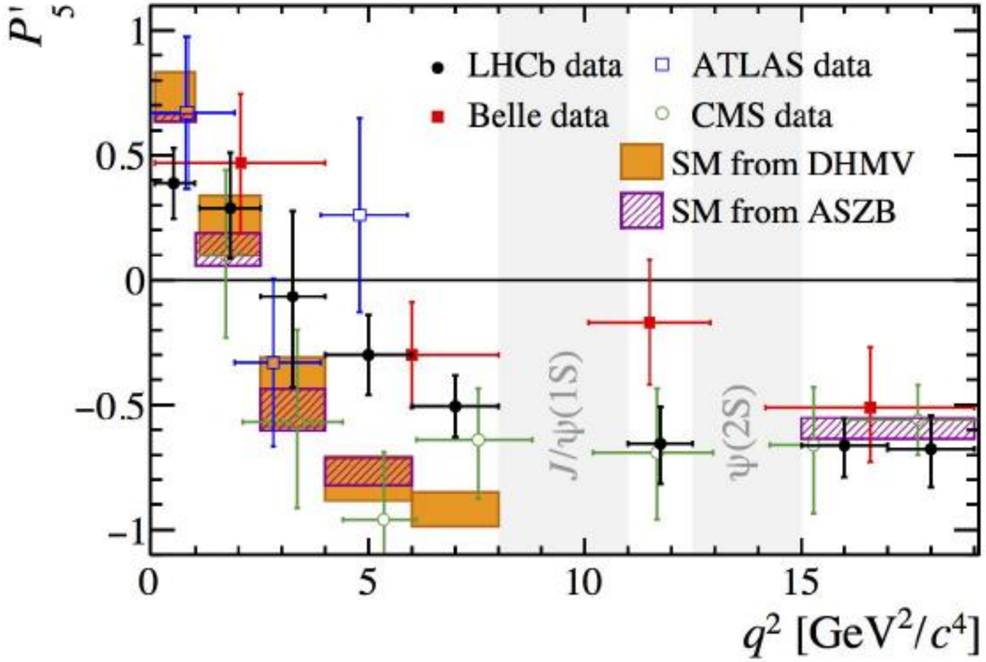
- **Flavor Anomaly can enter in global fits to help constraining NP parameters (Wilson coefficients):**





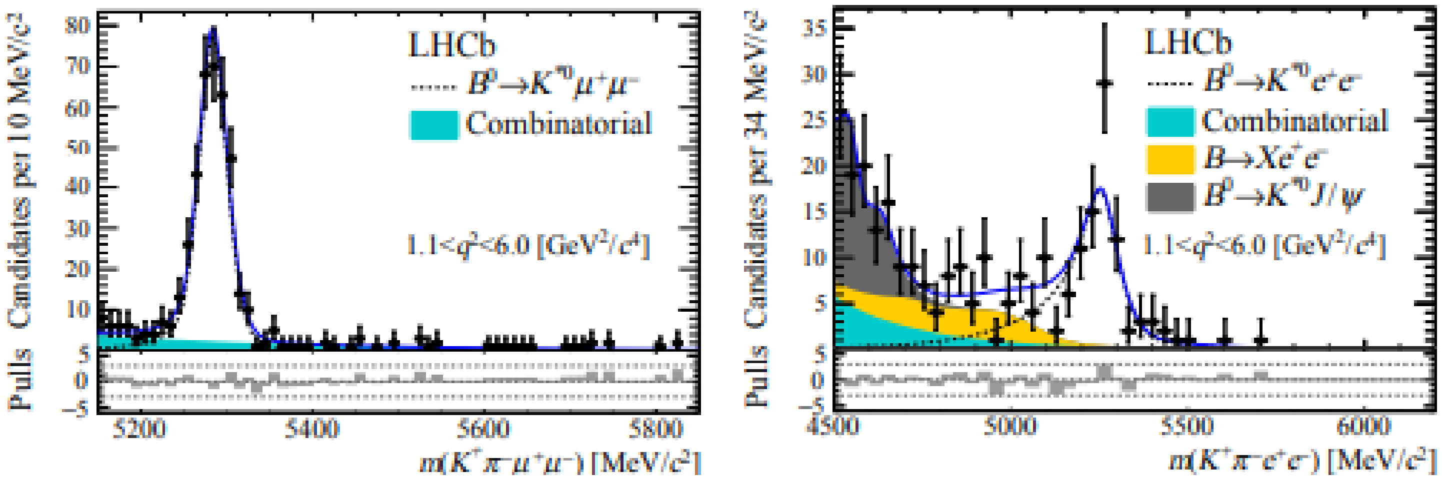
- **A significant flavor anomaly example:**
  - P'5 Angular Observable: a coefficients in the angular decay rate
    - Form factor uncertainties cancel at leading order
    - Significant tension of 3.4 sigma

$$\frac{d^4\Gamma}{d\cos\theta_\ell\, d\cos\theta_K\, d\phi\, dq^2} = \frac{9}{32\pi}\Big[J_{1s}\sin^2\theta_K + J_{1c}\cos^2\theta_K + (J_{2s}\sin^2\theta_K + J_{2c}\cos^2\theta_K)\cos 2\theta_l$$
$$+J_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi + J_4\sin 2\theta_K\sin 2\theta_l\cos\phi + J_5\sin 2\theta_K\sin\theta_l\cos\phi$$
$$+(J_{6s}\sin^2\theta_K + J_{6c}\cos^2\theta_K)\cos\theta_l + J_7\sin 2\theta_K\sin\theta_l\sin\phi$$
$$+J_8\sin 2\theta_K\sin 2\theta_l\sin\phi + J_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\Big]$$



$$P'_5 = \frac{J_5}{2\sqrt{-J_{2s}J_{2c}}}$$

- **As a test for the SM:**

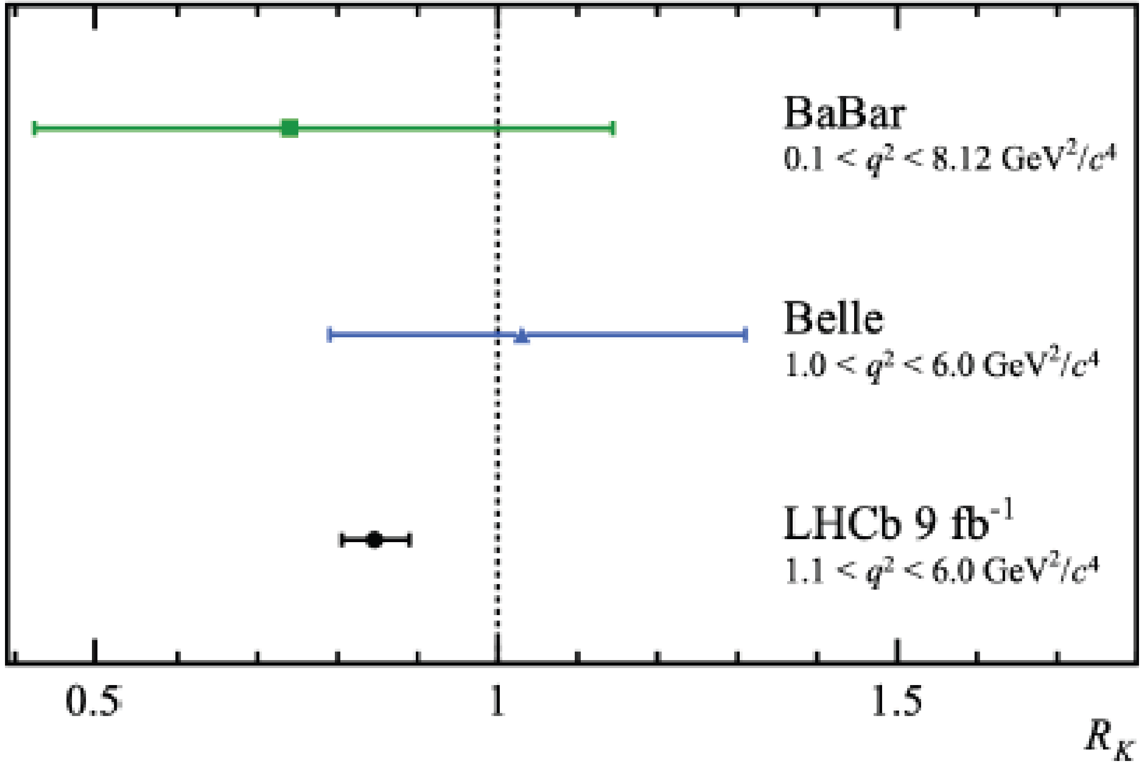


  - **Muon vs Electron:**
    - **Muon: very clean**
    - **Electron: difficult**
      - **ECAL: low efficiency/resolution**
      - **Bremsstrahlung affects resolution & efficiencies**
      - **Higher Trigger Threshold for e-**
    - **Many background sources & low statistics**

- **Control uncertainties by measuring double ratios:**

$$R_X \equiv \frac{\mathcal{B}(B \to X\,\mu\mu)}{\mathcal{B}(B \to X\,J/\psi\,(\to \mu\mu))} \frac{\mathcal{B}(B \to X\,J/\psi\,(\to ee))}{\mathcal{B}(B \to X\,ee)} = 1_{(SM)}$$

- **All LHCb measurements are below 1.**



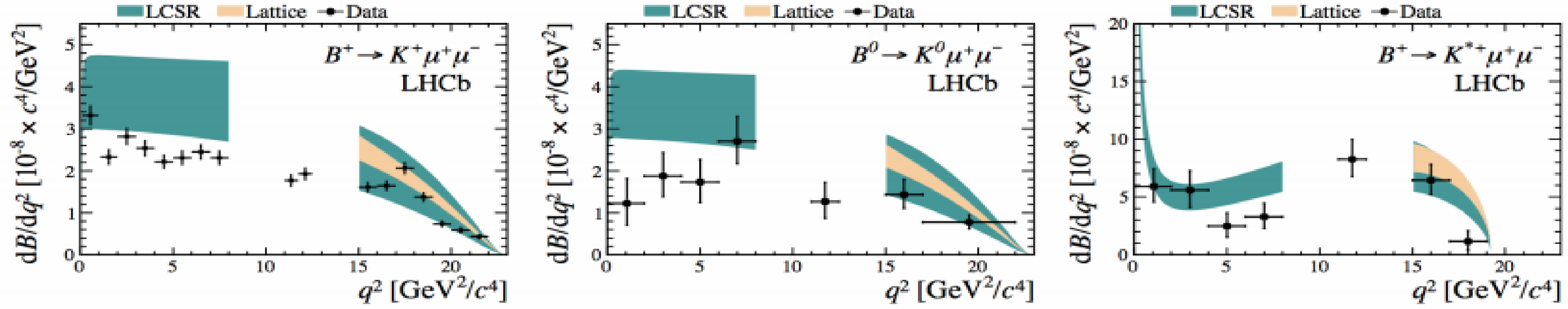
The most precise measurement:

$$R_K = \frac{\Gamma(B^+ \to K^+\mu^+\mu^-)}{\Gamma(B^+ \to K^+e^+e^-)} = 0.846^{+0.042+0.013}_{-0.039-0.012}$$

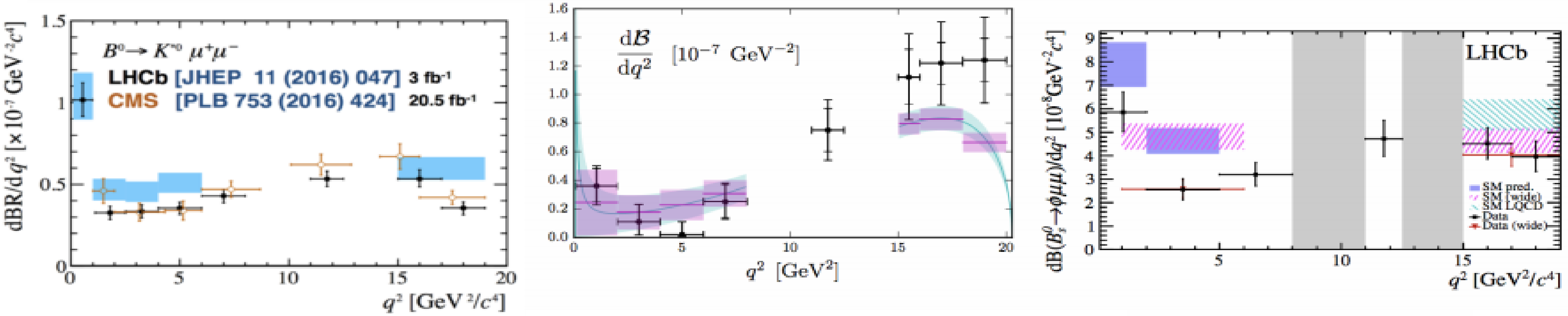
Deviation is 3.1 sigma.

b→sμμ decay rates (BF)

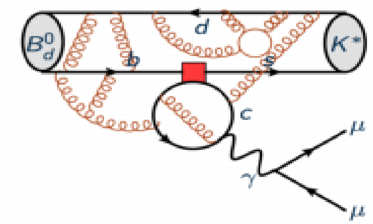
[JHEP06(2014)133]



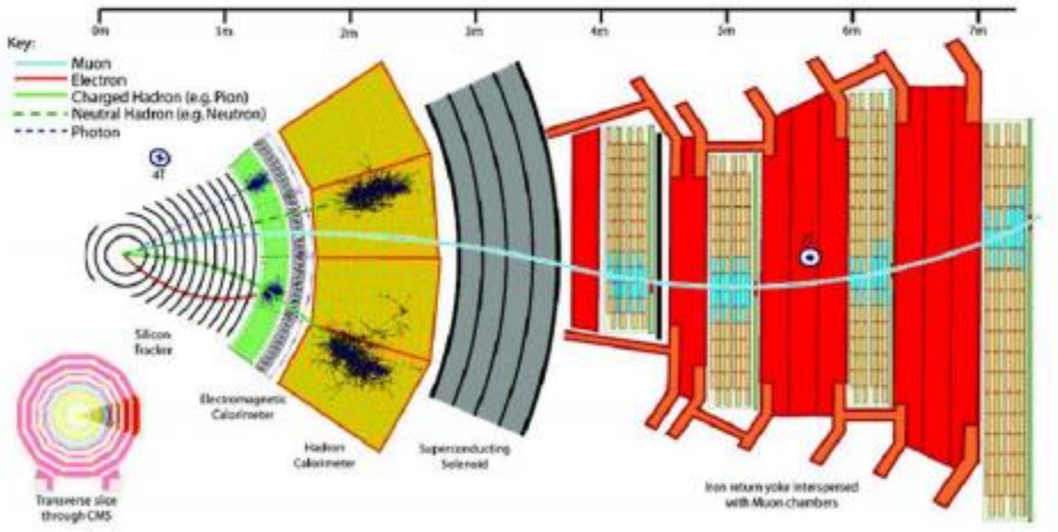
$B^0 \to K^{*0}\mu^+\mu^-$ [JHEP11(2016)047], $\Lambda_b \to \Lambda\mu^+\mu^-$ [JHEP06(2015)115] $B_s \to \phi\mu^+\mu^-$ [JHEP09(2015)179]



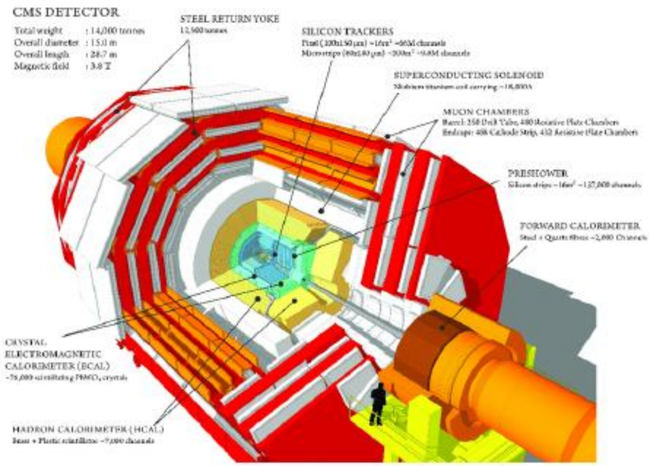
measurements tend to appear below theory, at low q2
SM predictions affected by large hadronic uncertainties



- **Relevant CMS elements:**
  - **Silicon tracker:** closest part to the collision point
    - Precise and efficient measurement of the trajectories of charged particles which curve in the presence of the magnetic field
    - Calculation of pT = qRB.
  - **Muon chamber:** outermost region of the detector
    - Muon identification with high efficiency



  - Collision delivered at tens of MHz we cannot save all of them
  - **two-level trigger system:**
    - L1: hardware processors select the acceptable events, using local information at the rate of around 90 kHz.
    - HLT: decreases the event rate to less than 1 kHz before data storage. The tracker information is not available at L1, and it works on software, allowing more elaborate reconstruction.



Current models used:

- **RT component:**
  - *q2bin 0-3:* DoubleCB($x; \bar{x}, \sigma, \alpha 1, \alpha 2, n1, n2$)
  - *q2bin 4-6:* CB($x; \bar{x}, \sigma 1, \alpha 1, n1$) + $f \times$CB($x; \bar{x}, \sigma 2, \alpha 2, n2$)
  - *q2bin 7:* CB($x; \bar{x}, \sigma 1, \alpha 1, n1$) + $f \times$Gauss($x; \bar{x}, \sigma 2$)
- **WT component:**
  - DoubleCB($x; \bar{x}, \sigma, \alpha 1, \alpha 2, n1, n2$)
- **Background:**
  - Exp($x; \lambda$)

CRYSTAL BALL FUNCTION

Crystal ball function:

$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases}$$

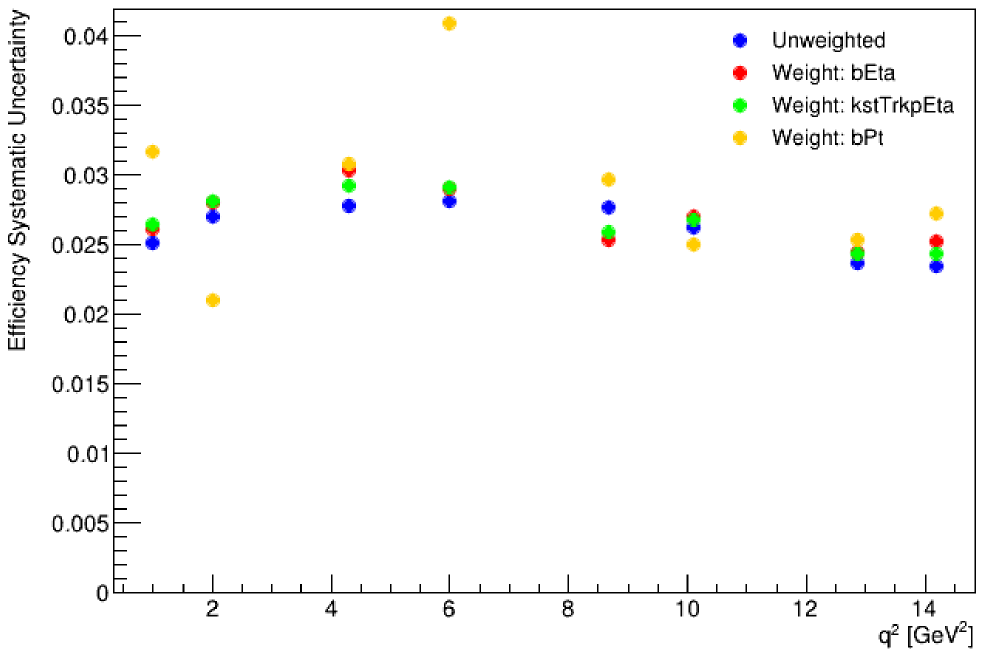
Double Crystal Ball:

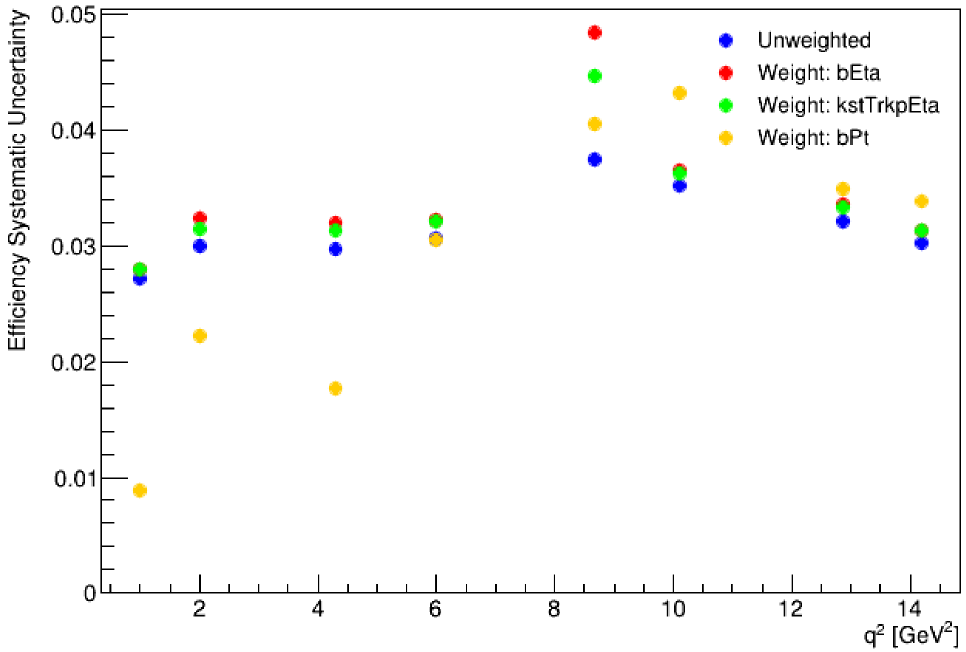
$$\text{DoubleCB}(x;\bar{x},\sigma,\alpha_1,\alpha_2,n_1,n_2) = \begin{cases} e^{-\frac{t^2}{2}}, & \text{if } -\alpha_1 < t < \alpha_2 \\ e^{-\frac{\alpha_1^2}{2}}\left[1 - \frac{\alpha_1}{n_1}(\alpha_1 + t)\right]^{-n_1}, & \text{if } t \le -\alpha_1 \\ e^{-\frac{\alpha_2^2}{2}}\left[1 - \frac{\alpha_2}{n_2}(\alpha_2 - t)\right]^{-n_2}, & \text{if } -t \ge \alpha_2 \end{cases}$$

$$t = \frac{x-\bar{x}}{\sigma}$$

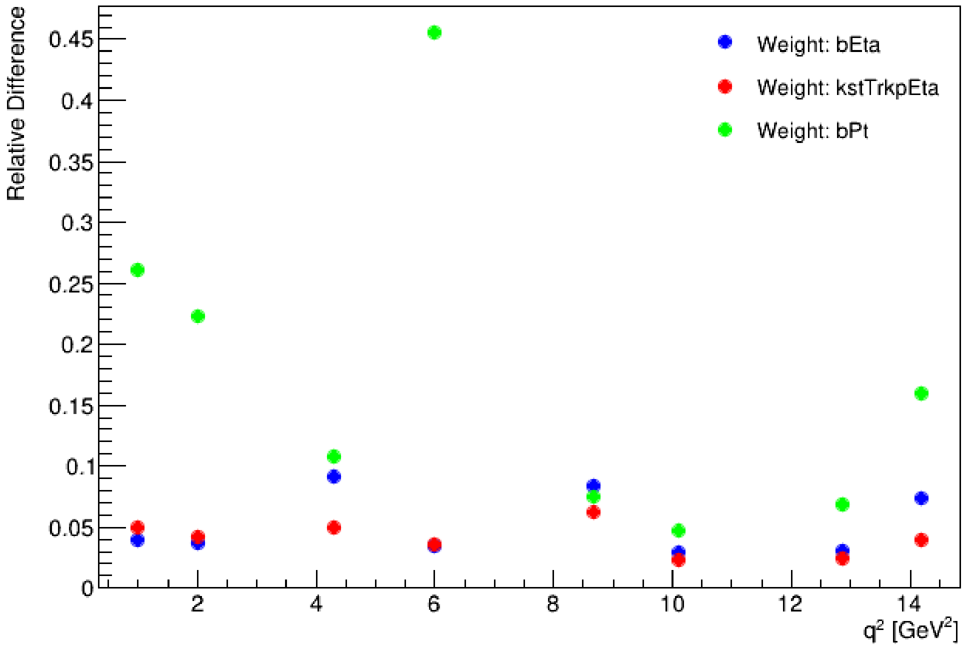
STEP TWO: EFFICIENCY

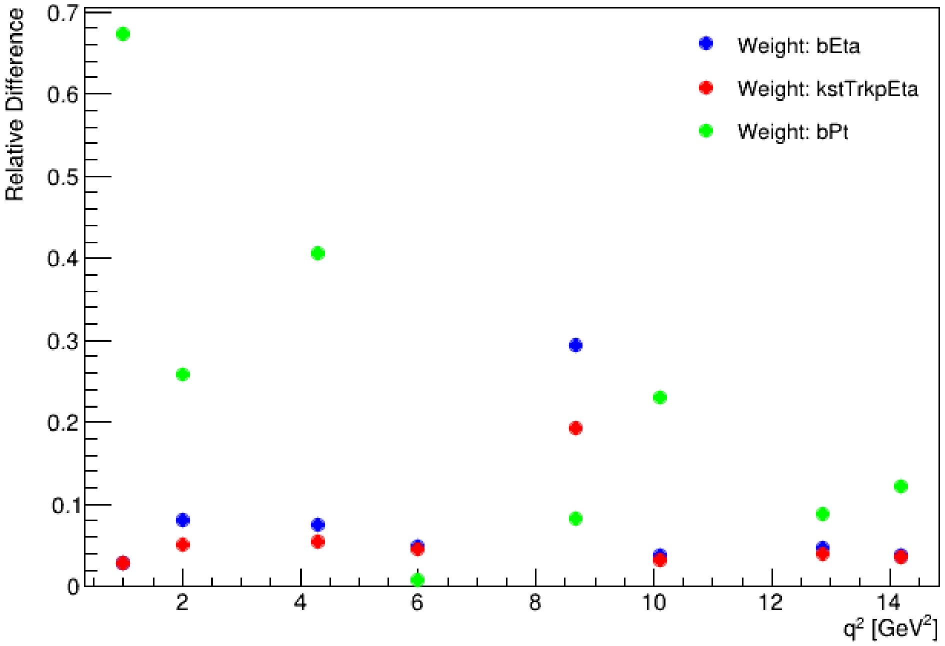
YEAR 2017 AND 2018 – EFFICIENCY RESULTS



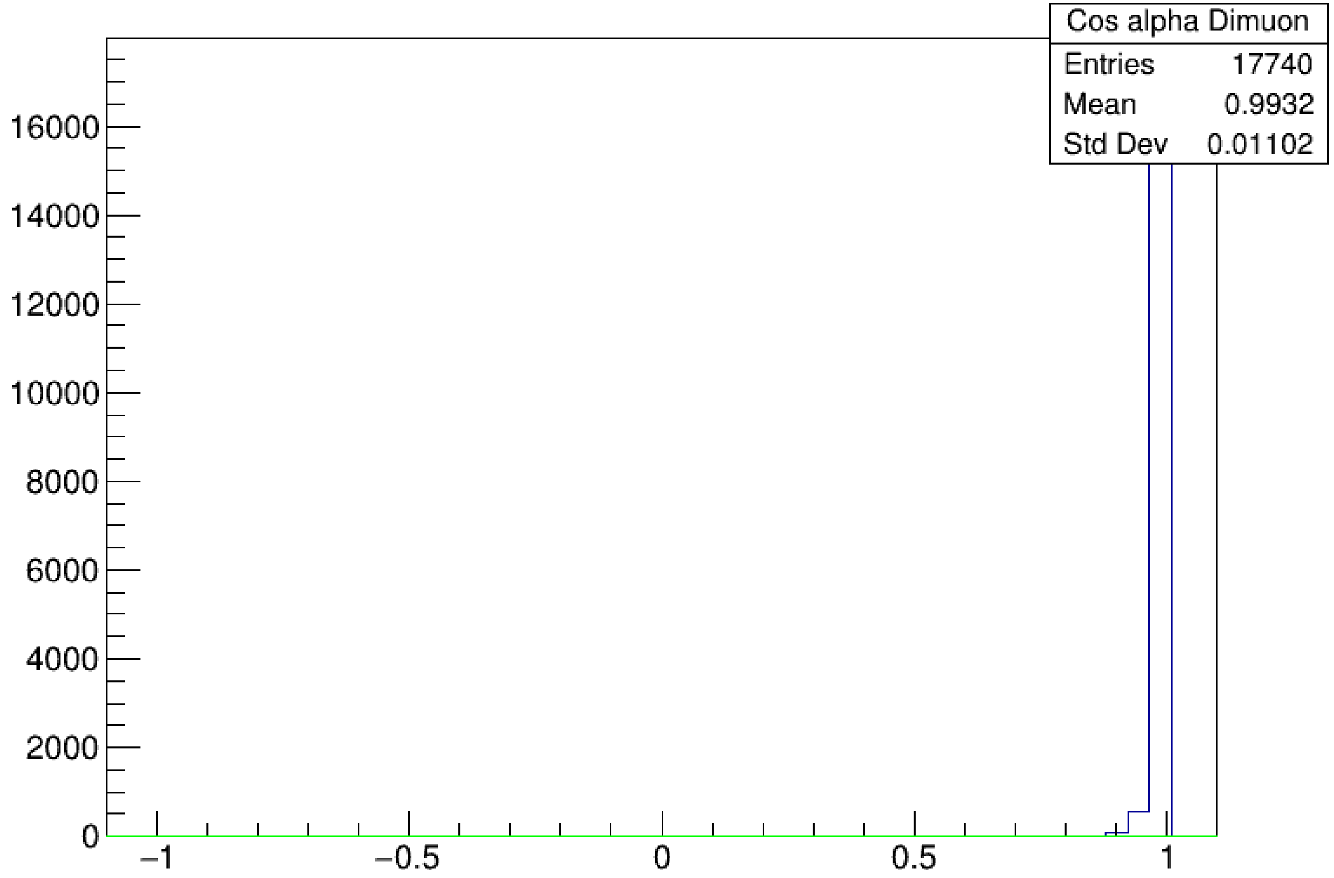


STEP TWO: EFFICIENCY
YEAR 2017 AND 2018 – RELATIVE DIFFERENCE RESULTS

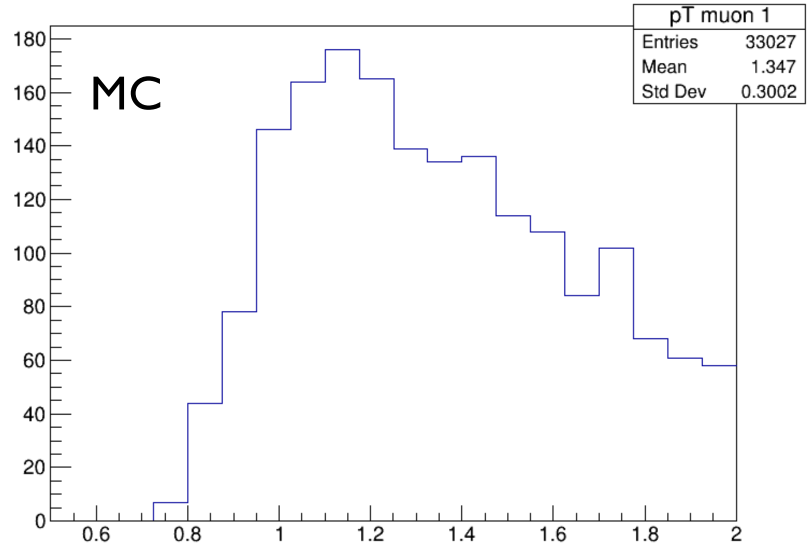


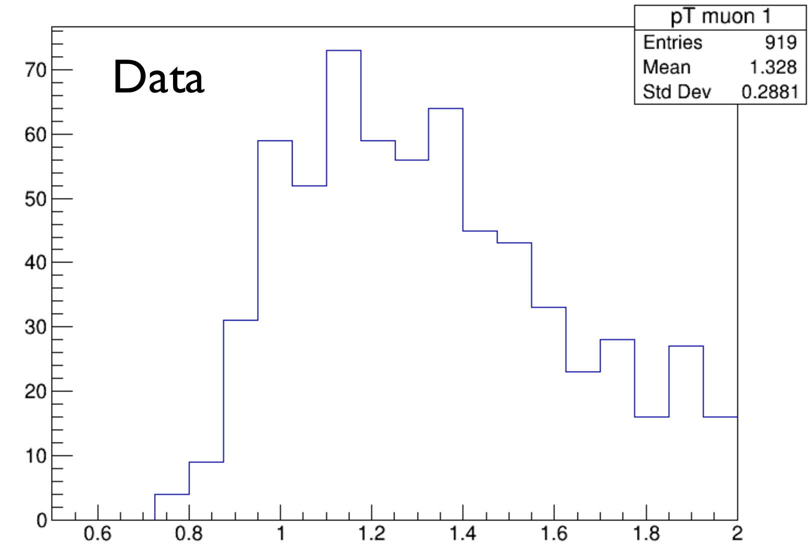


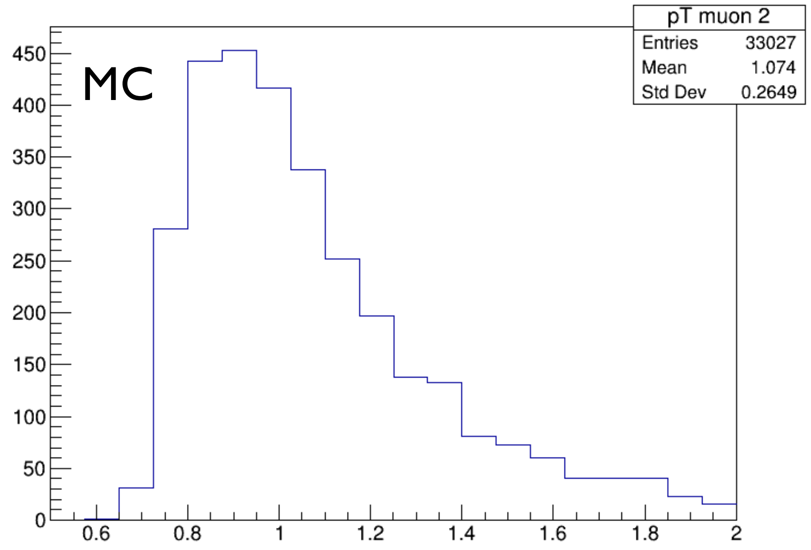
DISTRIBUTION OF COS (ALPHA)

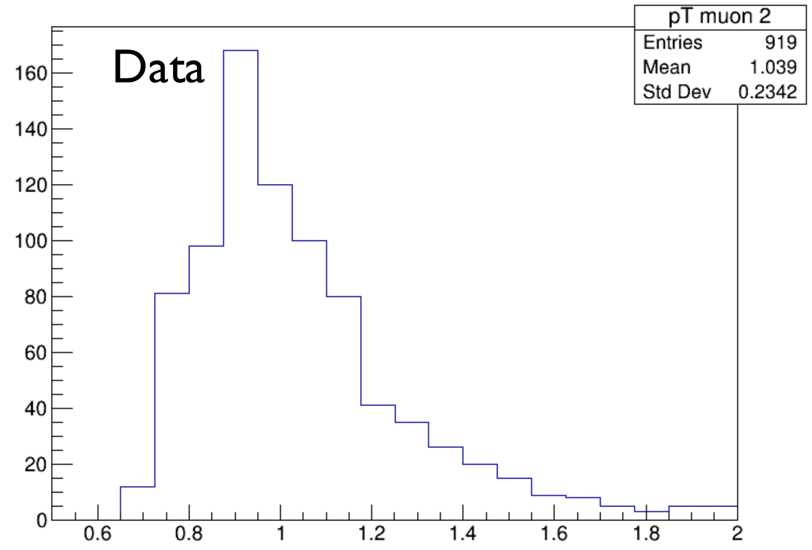


PT OF THE PARTICLES IN THE SYSTEM: MUONS

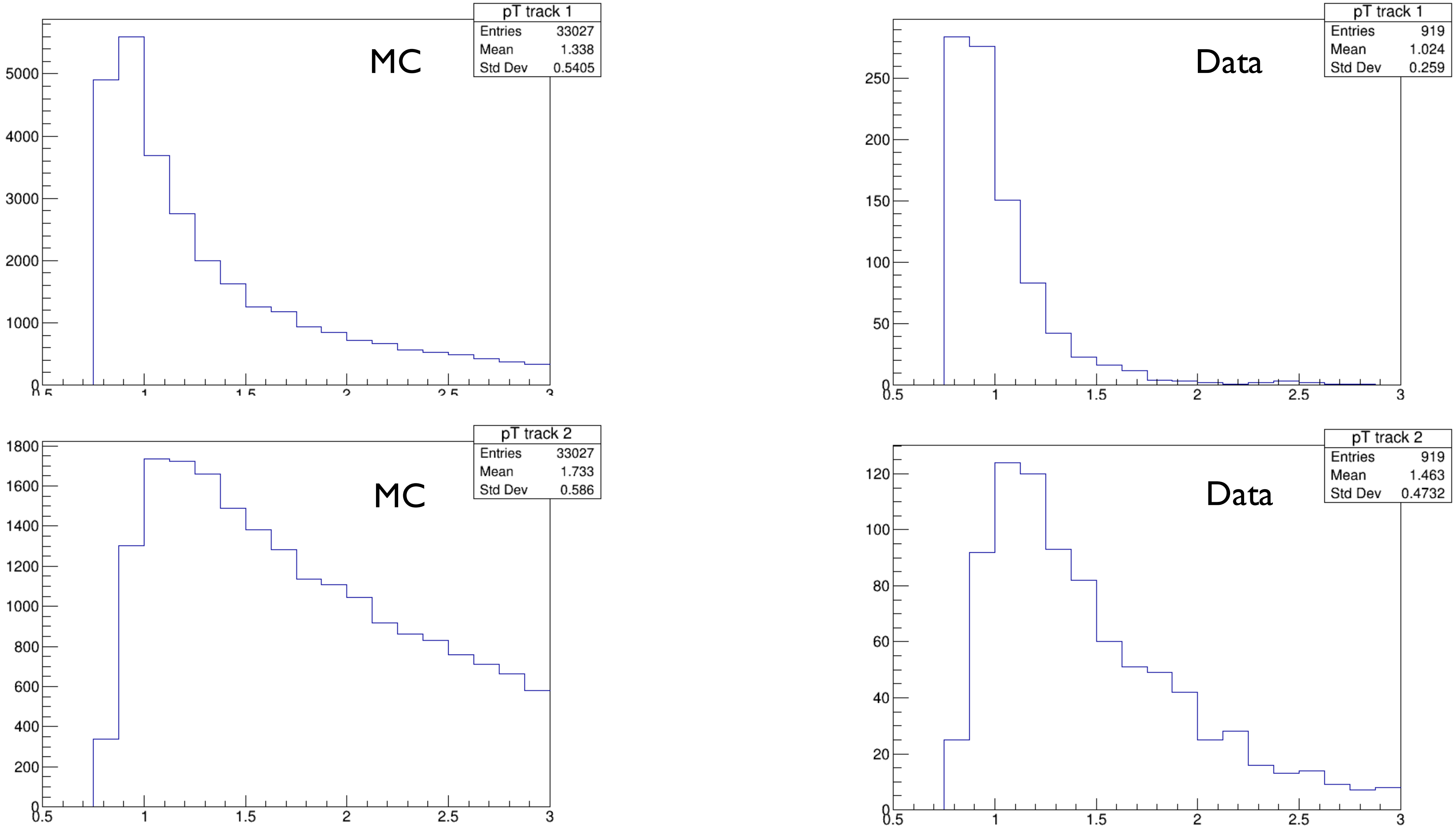




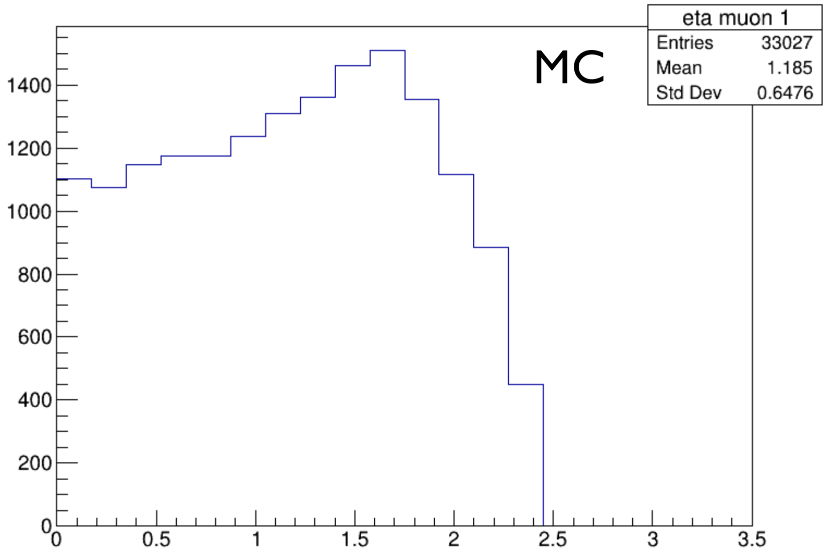


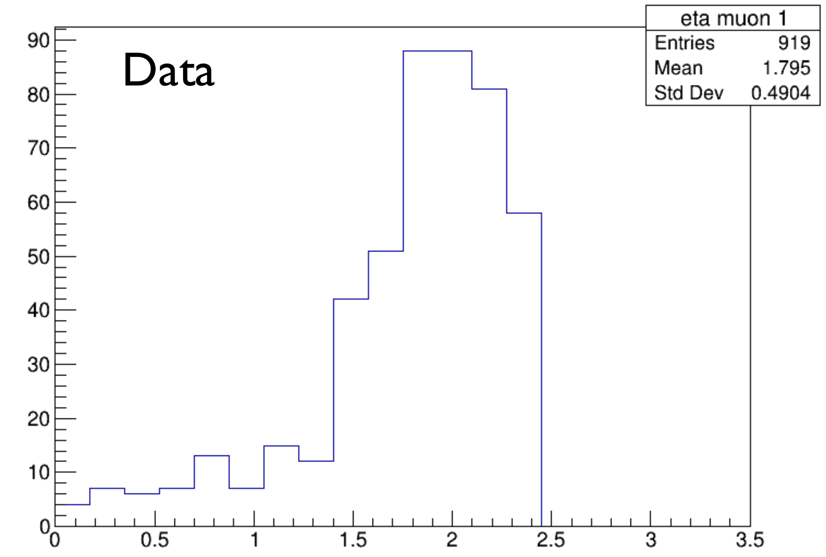


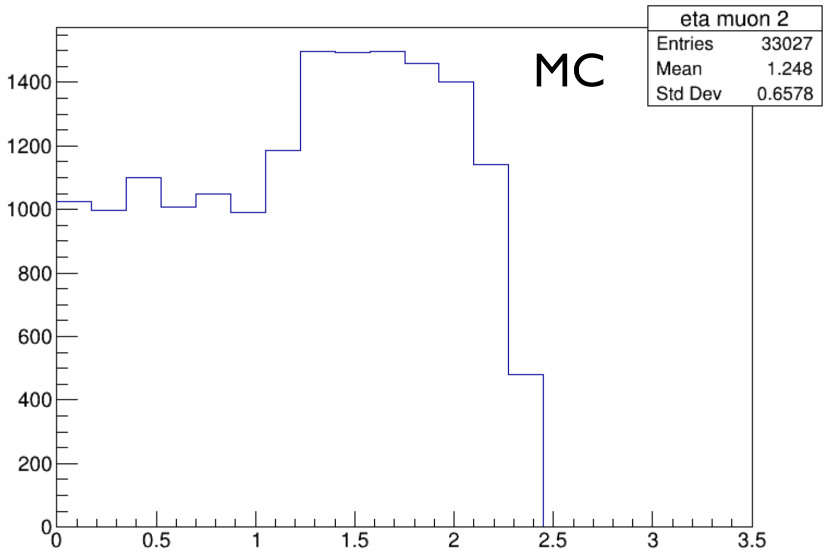
PT OF THE PARTICLES IN THE SYSTEM: TRACK 1 AND 2 (K AND PION)

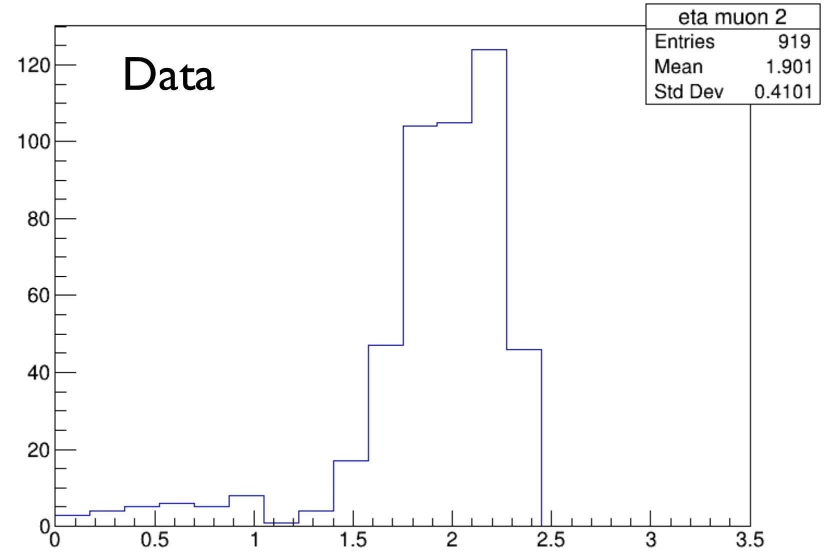


ETA OF THE PARTICLES IN THE SYSTEM: MUONS









Eta of the particles in the system: K and pion