Search for the lepton-flavour violating decays $B^+ \to K^+ \mu^\pm e^\mp$

J. Albrecht[1], J. Bhom[2], M. Chrzaszcz[4], G.Meier[1], T .Mombächer[1], M. Pikies[2], F. Polci[5], S. Reichert[1], N. Serra[3]

[1] Dortmund , [2] Krakow, [3] Zürich, [4]CERN, [5]LPNHE

Tue meeting, April 30, 2019

Yellow Pages

⇛ Twiki:
https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/B2hemu

⇛ Time scale:

- WG approval 21^{st} Nov. 2018
- Unblinding 25^{th} Apr. 2019

⇛ Journal target: PRL

⇛ Contact authors:
J.Bhom (Krakow), T. Mombächer (Dortmund), F. Polci (LPNHE)

Motivation

$\Rrightarrow$ Good „old” anomalies: P_5', R_K, R_{K^*} can implicate the existence of $b \to s\mu e$.

$\Rrightarrow$ LQ, neutrino CP violation, Z': $\mathcal{B}(B \to Ke\mu) = 10^{-8} - 10^{-10}$.

$\Rrightarrow$ Particularly interesting in LQ: $\mathcal{B}(B \to Ke\mu) = \frac{1-R_K}{0.23} \times 10^{-8}$ should be accessible within analysis sensitivity.

$\Rrightarrow$ Current best UL are from Babar (90% CL) [PR D73 (2006) 092001]:

- $\mathcal{B}(B^+ \to K^+ e^+ \mu^-) < 9.1 \times 10^{-8}$
- $\mathcal{B}(B^+ \to K^+ e^- \mu^+) < 13 \times 10^{-8}$

Motivation

⇛ Good „old" anomalies: P_5', R_K, R_{K^*} can implicate the existence of $b \to s\mu e$.

⇛ LQ, neutrino CP violation, Z': $\mathcal{B}(B \to Ke\mu) = 10^{-8} - 10^{-10}$.

⇛ Particularly interesting in LQ: $\mathcal{B}(B \to Ke\mu) = \frac{1-R_K}{0.23} \times 10^{-8}$ should be accessible within analysis sensitivity.

⇛ Current best UL are from Babar (90% CL) [PR D73 (2006) 092001]:

- $\mathcal{B}(B^+ \to K^+ e^+ \mu^-) < 9.1 \times 10^{-8}$
- $\mathcal{B}(B^+ \to K^+ e^- \mu^+) < 13 \times 10^{-8}$

⇛ Recently Belle also put its 5 cents [PRD 98 (2018) 071101]:

- $\mathcal{B}(B \to K^* e^+ \mu^-) < 1.6 \times 10^{-7}$
- $\mathcal{B}(B \to K^* e^- \mu^+) < 1.2 \times 10^{-7}$

⇛ NP predictions: 1503.01084, 1504.07928, 1503.07099, 1507.01412

Analysis Strategy

⇛ Blind analysis: signal window looked at after finalizing analysis procedure.

⇛ Analysis strategy:

- Stripping
- Loose preselection
- Target vetos
- Hard MVA and PID selection
- Upper limit setting
- Book airplane ticket to Stockholm (optional ;))

⇛ Dataset: $3~\mathrm{fb}^{-1}$, Run1.

Normalization

⇛ Typically the $b\bar{b}$ cross section has large uncertainty.
⇛ It is more beneficial to normalize the decay rate to well know branching fraction ⇛ reduce systematics:

$$\mathcal{B}(B^+ \to K^+ \mu^\pm e^\mp) = N_{B^+ \to K^+ \mu^\pm e^\mp} \times \frac{\mathcal{B}(B^+ \to K^+ J/\psi(\mu^+\mu^-))}{\varepsilon_{B^+ \to K^+ \mu^\pm e^\mp}}$$

$$\frac{\varepsilon_{B^+ \to K^+ J/\psi(\mu^+\mu^-)}}{N_{B^+ \to K^+ J/\psi(\mu^+\mu^-)}} = N_{B^+ \to K^+ \mu^\pm e^\mp} \times \alpha$$

⇛ Some systematic cancel in the efficiency ratio.

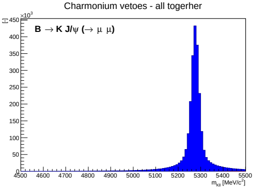
Stripping

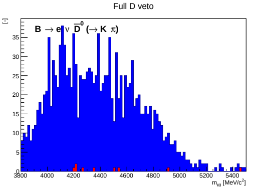
Particle or event	Variable	Cut
Event	n_{SPD}	< 600
B	$\|m - m_{\mathrm{PDG}}\|$	$< 1500\,\mathrm{MeV}$
	$\chi^2_{\mathrm{vtx}}/\mathrm{dof}$	< 9
	χ^2_{FD} wrt. PV	> 100
	χ^2_{IP} wrt. PV	< 25
	DIRA wrt. PV	> 0.9995
K	p_T	$> 400\,\mathrm{MeV}$
	χ^2_{IP}	> 9
e	p_T	$> 300\,\mathrm{MeV}$
	χ^2_{IP} wrt. PV	> 9
	$PIDe$	> 0
μ	p_T	$> 300\,\mathrm{MeV}$
	χ^2_{IP} wrt.PV	> 9
	ISMUON	True
	HASMUON	True
$e\mu$ pair	$m(e\mu)$	$> 100\,\mathrm{MeV}$
	$\chi^2_{\mathrm{vtx}}(e\mu)/\mathrm{dof}$	< 9
Dimuon	p_T	$> 0\,\mathrm{MeV}$
	m	$< 5500\,\mathrm{MeV}$
	$\chi^2_{\mathrm{vtx}}/\mathrm{dof}$	< 9
	χ^2_{FD} wrt. PV	> 16
	χ^2_{IP} wrt. PV	> 0
Dielectron	p_T	$> 0\,\mathrm{MeV}$
	m	$< 5500\,\mathrm{MeV}$
	$\chi^2_{\mathrm{vtx}}/\mathrm{dof}$	< 9
	χ^2_{FD} wrt. PV	> 16
	χ^2_{IP} wrt. PV	> 0

Preselection: veto peaking backgrounds

⇛ Double semileptonic decays:
$B \to e\nu D X$, $D \to \mu\nu K X$ or $B \to \mu\nu D X$, $D \to e\nu K X$
Efficiently removed by: $m_{K\ell} > 1885\,\mathrm{MeV}/c^2$
⇛ Charmonium decays:
$B \to J/\psi/\psi(2S)K^+$, with missID daughters.

mass swap	mass region vetoed (MeV)
K with μ mass	$3000 < m_{K^-\mu^+} < 3200$ $3630 < m_{K^-\mu^+} < 3740$
e with μ mass	$2950 < m_{e^-\mu^+} < 3200$ $3630 < m_{e^-\mu^+} < 3740$
K with e mass	$3000 < m_{K^+e^-} < 3200$ $3630 < m_{K^+e^-} < 3740$
μ with e mass	$3000 < m_{\mu^+e^-} < 3200$ $3630 < m_{\mu^+e^-} < 3740$



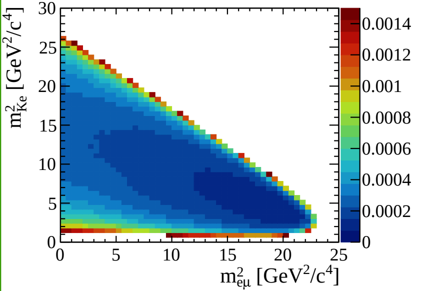


Signal/Control Channel Model

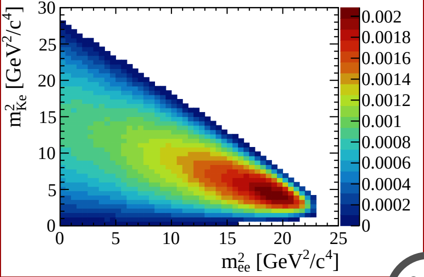
$\Rrightarrow$ The MC models that have been used in the analysis needs to be updated:

- $B^+ \to K^+ e^+ e^-$ decay was generated with PHSP model. Needs to be BTOSBALL.
- $B^+ \to K^+ \mu^\pm e^\mp$ decay was generated with BTOSBALL model. Needs to be PHSP.

Signal mode



Control mode



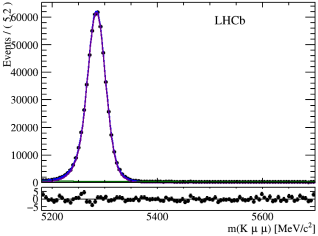
Trigger lines

Channel	Particle	L0 trigger
$B \to Ke\mu$	μ	L0MuonDecision
$B \to KJ/\psi(\to \mu\mu)$	$\mu^{\pm}$	L0MuonDecision
$B \to KJ/\psi(\to ee)$	$e^{\pm}$	L0ElectronDecision

Channel	Hlt1 trigger	Hlt2 trigger
$B^{+} \to K^{+}\mu^{\pm}e^{\mp}$	TrackMuonDecision TrackAllL0Decision	TopoMu[2,3]BodyBBDTDecision Topo[2,3]BodyBBDTDecision SingleMuonDecision SingleMuonLowPTDecision
$B^{+} \to K^{+}J/\psi(\mu^{+}\mu^{-})$	TrackMuonDecision TrackAllL0Decision	TopoMu[2,3]BodyBBDTDecision Topo[2,3]BodyBBDTDecision SingleMuonDecision SingleMuonLowPTDecision
$B^{+} \to K^{+}J/\psi(e^{+}e^{-})$	TrackAllL0Decision TopoE[2,3]BodyBBDTDecision	Topo[2,3]BodyBBDTDecision

Data/MC differences

⇛ MC reweighted to correct for data/simulation differences.
⇛ Weights extracted from binned distributions of nTracks, $p_T(B)$ and Vtx χ^2 from $B^+ \to K^+ J/\psi(\mu^+\mu^-)$.

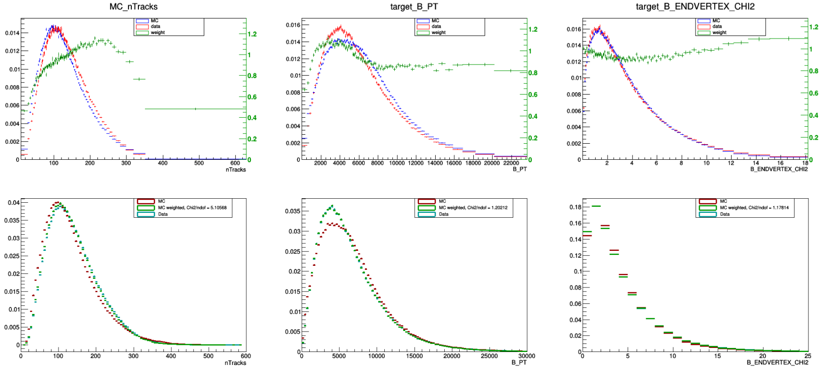


⇛ Splotted data $B^+ \to K^+ J/\psi(\mu^+\mu^-)$: double Crystal-Ball and exp function.

Data/MC differences

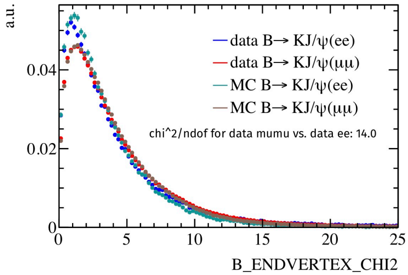
⇛ Iterative procedure to correct one variable at a time.

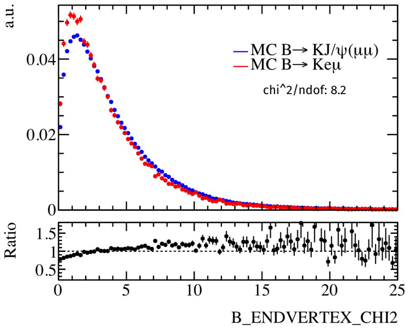
⇛ Convergence after first iteration of variables.



Electron - Muon difference

⇛ The weights are determined from the muon mode.
⇛ The question is the VTX χ^2 the same for electrons?





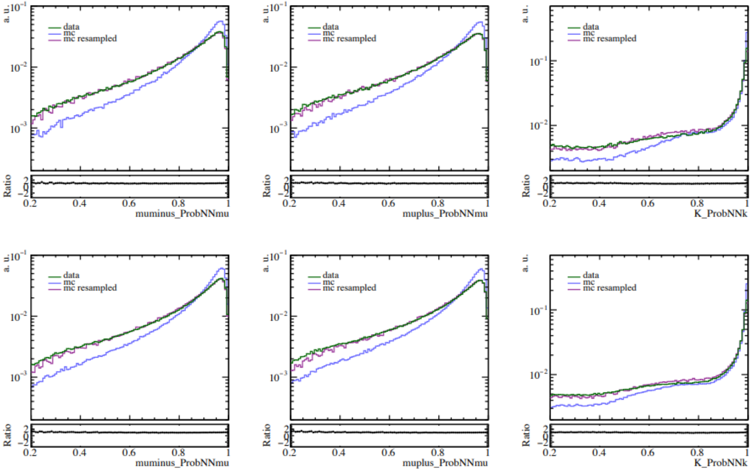
⇛ Assign 1.4% as systematic as change of normalisation constant.

PID Resampling

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

2011

2012

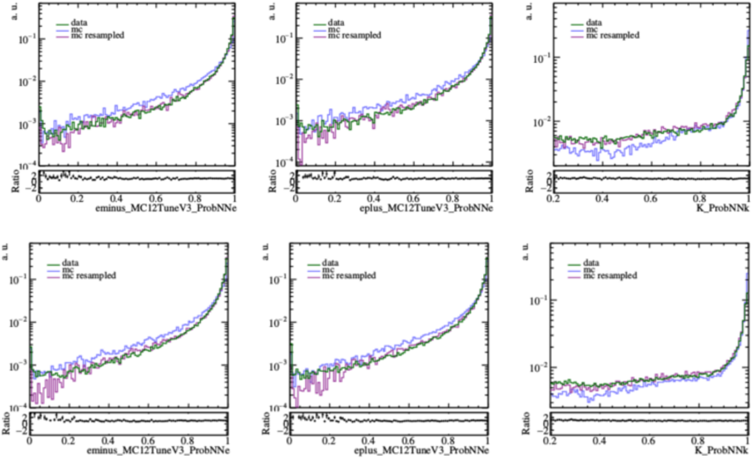


PID Resampling

$B^+ \to K^+ J/\psi(\to e^+e^-)$

2011

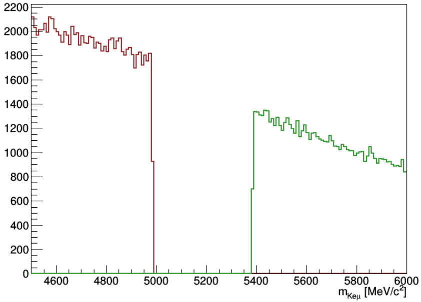
2012



BDT strategy

⇛ Combinatorial background suppression via BDT using upper sideband as background proxy

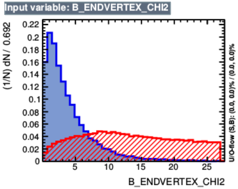
⇛ Partially reconstructed background rejected via BDTHOP using lower sideband as background proxy

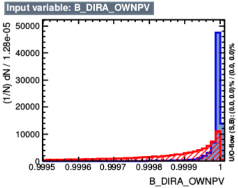


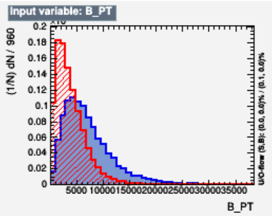
⇛ k-Folding technique used, with $k = 10$ folds.

BDT training and results

- the transverse momentum p_T of the B candidate,
- the momentum p of the B candidate,
- the impact parameter χ^2, χ^2_{IP}, of the B candidate,
- the direction angle (DIRA) of the B candidate,
- the quality of the $Ke\mu$ vertex χ^2,
- the B flight distance χ^2
- the impact parameter χ^2, χ^2_{IP}, of the kaon
- the minimum and maximum of electron and muon IP candidates
- the cut based isolation variables from $B_s \to \mu\mu$ analysis: `B_relinfo_BSMUMUTRACKPLUSISOTWO_L1,2` and `B_relinfo_cone_pt_asym_H`.

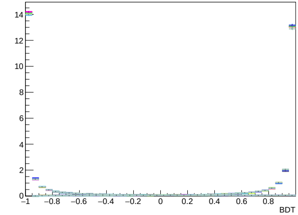


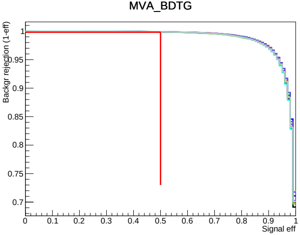




BDT training and results

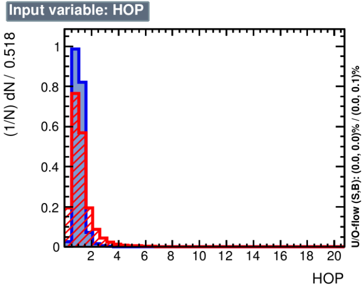
- the transverse momentum p_T of the B candidate,
- the momentum p of the B candidate,
- the impact parameter χ^2, χ^2_{IP}, of the B candidate,
- the direction angle (DIRA) of the B candidate,
- the quality of the $Ke\mu$ vertex χ^2,
- the B flight distance χ^2
- the impact parameter χ^2, χ^2_{IP}, of the kaon
- the minimum and maximum of electron and muon IP candidates
- the cut based isolation variables from $B_s \rightarrow \mu\mu$ analysis: `B_relinfo_BSMUMUTRACKPLUSISOTWO_L1,2` and `B_relinfo_cone_pt_asym_H`.





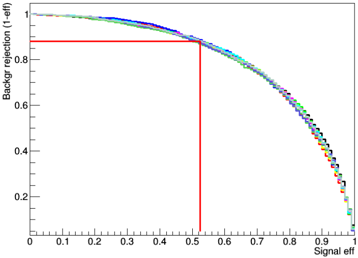
BDTHOP strategy

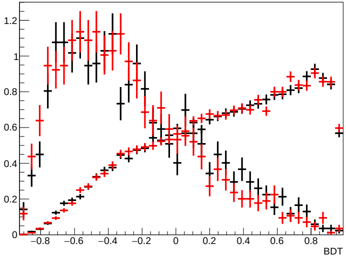
⇛ Apply a cut: $BDT > 0.98$ prior to training
⇛ Use same inputs and strategy as for BDT with additional input of HOP
⇛ Reminder: background proxy is lower mass sideband.



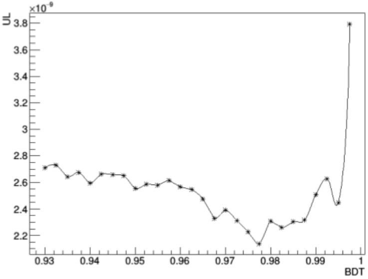
BDTHOP results

⇛ Clearly discriminating power remains.

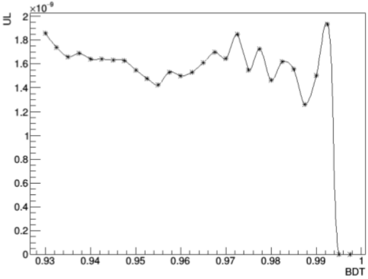




- ⇛ Standard question in RD: How to optimize the selection?
- ⇛ We used $\mathrm{CL_s}$ method for optimization [Read:451614].
- ⇛ What dataset do you use?
- ⇛ We split the datasets for $B^+ \to K^+\mu^-e^+$ and $B^+ \to K^+\mu^+e^-$ and optimize separately.



(a) $B^+ \to K^+\mu^+e^-$



(b) $B^+ \to K^+\mu^-e^+$

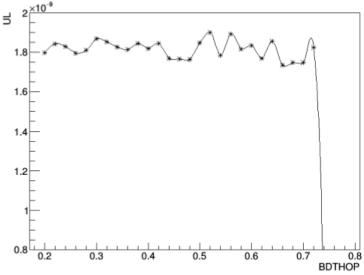
Optimization

⇛ Standard question in RD: How to optimize the selection?

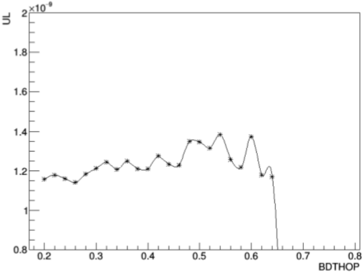
⇛ We used $\mathrm{CL_s}$ method for optimization [Read:451614].

⇛ What dataset do you use?

⇛ We split the datasets for $B^+ \to K^+\mu^-e^+$ and $B^+ \to K^+\mu^+e^-$ and optimize separately.



(a) $B^+ \to K^+\mu^+e^-$

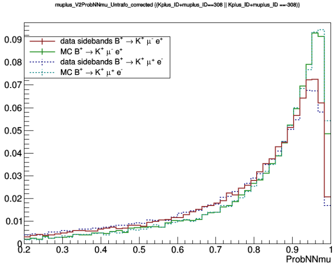


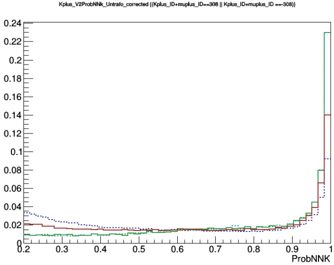
(b) $B^+ \to K^+\mu^-e^+$

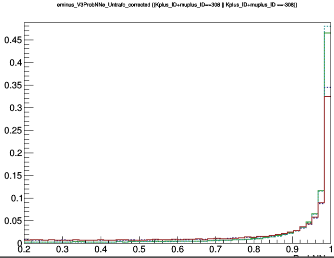
Optimization - PID

⇛ After the BDT and BDTHOP cuts there isn't much events left to perform PID optimization.

⇛ Decided to put a conservative cuts.





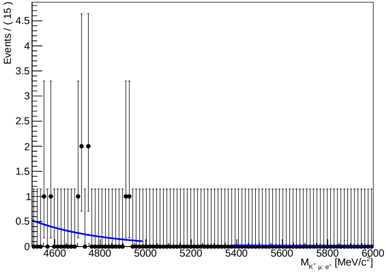


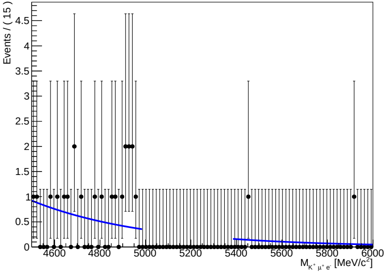
⇛ Conservative cuts:

$$\mathrm{ProbNNmu} > 0.7\,, \quad (1)$$
$$\mathrm{ProbNNe} > 0.65\,, \quad (2)$$
$$\mathrm{ProbNNk} > 0.65.$$

Optimization summary

⇛ After the full selection this is how our blinded data set looks like:





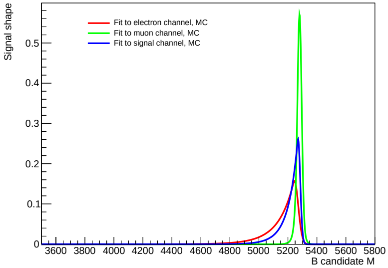
Mode	Expected background
$B^+ \to K^+\mu^+e^-$	3.94 ± 1.15
$B^+ \to K^+\mu^-e^+$	0.88 ± 0.64

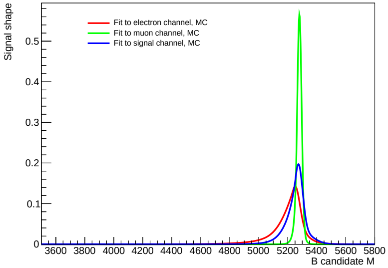
Signal Mass Model

⇛ We need to know the signal model.

⇛ We used a data-driven procedure to correct the parameters P:

$$P_{\mu e}^{\text{pred}} = P_{ee}^{\text{data}} + (P_{\mu e}^{\text{MC}} - P_{ee}^{\text{MC}}) \cdot \frac{P_{ee}^{\text{data}} - P_{\mu\mu}^{\text{data}}}{P_{ee}^{\text{MC}} - P_{\mu\mu}^{\text{MC}}}. \quad (4)$$



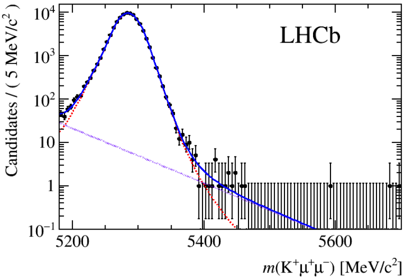


⇛ Cross-checked with the $B \to e\mu$ analysis method

Normalization

⇛ The normalization factor:

$$\alpha = \frac{\mathcal{B}(B^+ \to K^+ J/\psi(\mu^+\mu^-))}{\varepsilon_{B^+\to K^+\mu^\pm e^\mp}} \frac{\varepsilon_{B^+\to K^+ J/\psi(\mu^+\mu^-)}}{N_{B^+\to K^+ J/\psi(\mu^+\mu^-)}}$$



$\mathcal{B}(B^+ \to K^+ J/\psi(\mu^+\mu^-)) = (6.021 \pm 0.174) \times 10^{-5}$ PDG

⇛ Control channel yields:

- 2011:
  $N_{B^+\to K^+ J/\psi(\mu^+\mu^-)} = 26940 \pm 170$
- 2012:
  $N_{B^+\to K^+ J/\psi(\mu^+\mu^-)} = 59220 \pm 250$

⇛ Combined alpha:

Decay	$\alpha/10^{-9}$
$B^+ \to K^+\mu^- e^+$	1.97 ± 0.18
$B^+ \to K^+\mu^+ e^-$	2.21 ± 0.19

Peaking backgrounds

Monte Carlo samples	full mass region	signal region
$B^+ \to K^+\mu^+\mu^-$	0.12 ± 0.05	0.10 ± 0.04
$B^+ \to K^+e^+e^-$	0.0080 ± 0.0071	0.0068 ± 0.0060
$B^+ \to K^+J/\psi(\to \mu\mu)$	< 0.53	< 0.053
$B^+ \to K^+J/\psi(\to e^+e^-)$	< 1.05	< 0.21
$B^0 \to K^{*0}e^+e^-$	< 0.0014	< 0.00014
$\Lambda_b \to pK^-\mu\mu$	0.0072 ± 0.0030	0.0029 ± 0.0014
$\Lambda_b \to pK^-e^+e^-$	< 0.0012	< 0.00048
$\Lambda_b \to pK^-J/\psi(\to \mu\mu)$	< 0.26	< 0.013
$\Lambda_b \to pK^-J/\psi(\to e^+e^-)$	< 1.08	< 0.054
$B^- \to D(\to K^-\mu\nu)\nu\mu$	< 2.5	< 0.050
$B^+ \to D(\to K^+e^-\nu)e^+\mu$	< 0.50	< 0.010
$B^+ \to D(\to K^+\pi^-)e^+\nu$	< 2.8	< 0.056
$B^+ \to K^+\pi^+\pi^-$	1.8 ± 3.2	0.049 ± 0.82

$\Rrightarrow$ All backgrounds are suppressed to a negligible level.

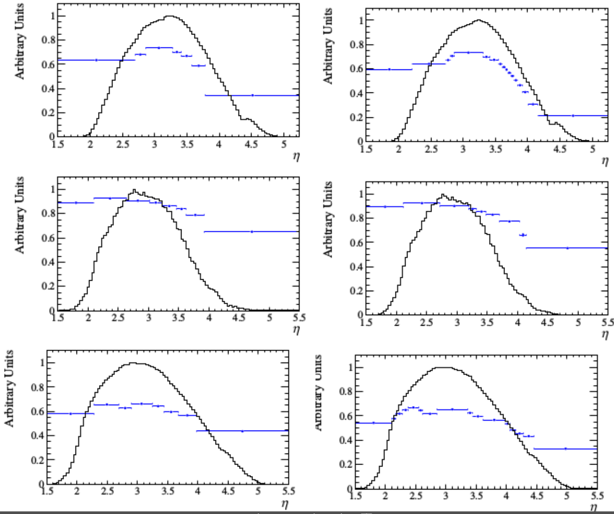
Systematics

Effect	$B^+ \to K^+\mu^+e^-$	$B^+ \to K^+\mu^-e^+$
Data-MC corrections	1.0	1.0
Electron-muon differences	1.4	1.4
MC model weights	0.2	0.2
Fitting model	2.1	2.1
PID resampling - binning	3.3	4.6
PID resampling - sWeighting	3	3
Background (not in %)	0.68	1.67
Trigger	1.0	1.0
Normalisation	6.8	6.6
Total	8.6	9.1

Systematics, PID - binning

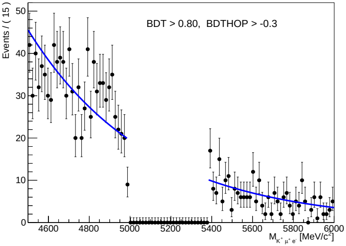
⇛ To perform resampling we bin the PID efficiency in nTracks, p and η.

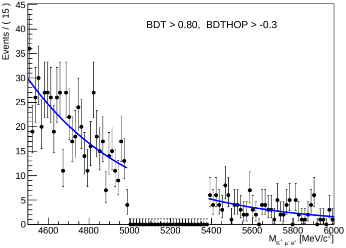
⇛ To access the systematics the finner and coarser binnings are applied.



Systematics, Background model

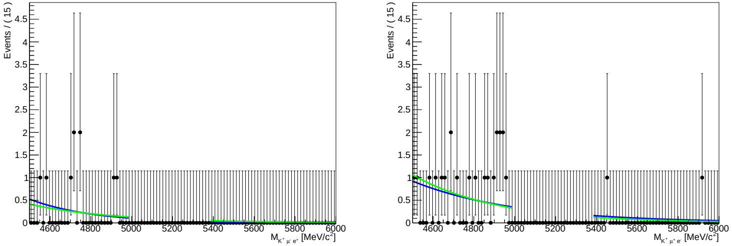
⇛ In the nominal fit we assume exponential shape of the background.
⇛ The alternative model is determined with a loose selection:





Systematics, Background model

⇛ Now compared to previous fit:

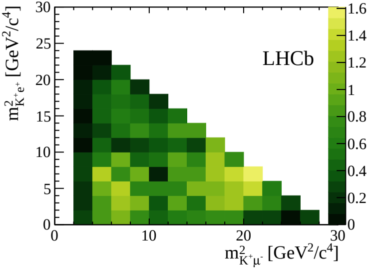


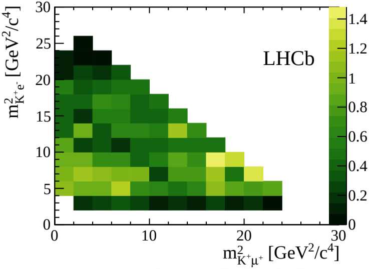
Channel	Background systematic	Nominal fit	Alternative fit
$B^+ \to K^+\mu^+e^-$	0.60	3.93 ± 1.14	3.33 ± 0.69
$B^+ \to K^+\mu^-e^+$	0.43	0.88 ± 0.63	1.30 ± 0.43

Systematics, Efficiency maps

⇛ The upper limits are set assuming PHSP model.

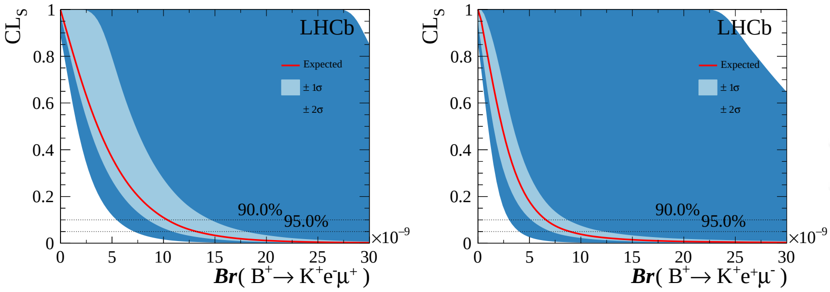
⇛ We will also provide the efficiency maps so theorists can interpret the results in their favorite model:





Upper limits

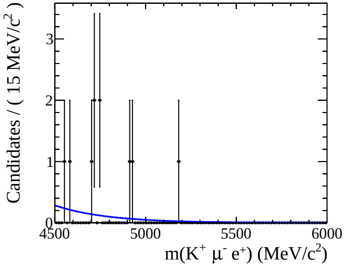
⇛ Upper limits set with $\mathrm{CL_s}$ method in GammaCombo.
⇛ This version of $\mathrm{CL_s}$ takes into account the signal and background shape information.
⇛ This gains the better expected upper limits by 25%. wrt. counting method

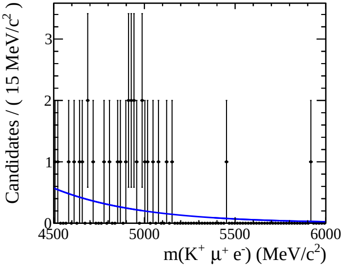


Unblinded results

⇛ On the 25^{th} April we have unblinded our dataset.

Channel	Expected background events	Observed
$B^+ \to K^+\mu^+e^-$	3.93 ± 1.14	2
$B^+ \to K^+\mu^-e^+$	0.88 ± 0.63	1

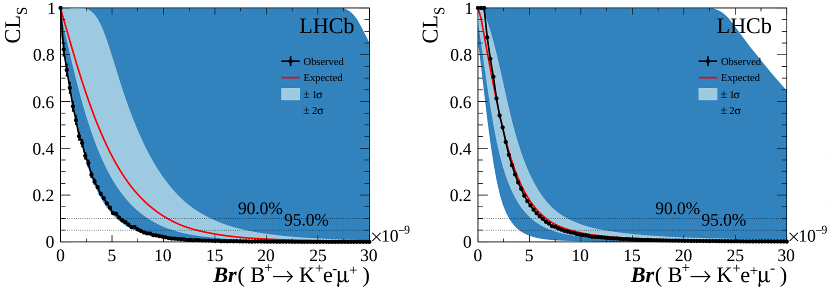




Unblinded results

⇛ On the 25^{th} April we have unblinded our dataset.

Channel	Expected background events	Observed
$B^+ \to K^+\mu^+e^-$	3.93 ± 1.14	2
$B^+ \to K^+\mu^-e^+$	0.88 ± 0.63	1



Unblinded results

$\Rrightarrow$ On the 25^{th} April we have unblinded our dataset.

Channel	Expected background events	Observed
$B^+ \to K^+\mu^+e^-$	3.93 ± 1.14	2
$B^+ \to K^+\mu^-e^+$	0.88 ± 0.63	1

$\Rrightarrow$ At 90% (95%) confidence level, the observed upper limits for the branching fractions are found to be

$$\mathcal{B}(B^+ \to K^+e^+\mu^-) < 6.3(8.2) \times 10^{-9},$$

$$\mathcal{B}(B^+ \to K^+e^-\mu^+) < 5.7(7.6) \times 10^{-9}.$$

$$\mathcal{B}(B^+ \to K^+e^\pm\mu^\mp) < 5.7(7.8) \times 10^{-9}.$$

Summary

- $\Rrightarrow$ Search for $B^+ \to K^+\mu^\pm e^\mp$decays has been performed
- $\Rrightarrow$ Over a order of magnitude improvement wrt. to BaBar results
- $\Rrightarrow$ We target PRL.
- $\Rrightarrow$ No significant excess of events has been observed
- $\Rrightarrow$ Proponents ask the collaboration to approve this analysis.

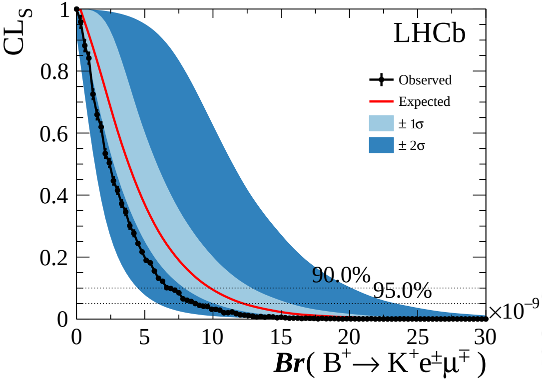
Summary

- ⇛ Search for $B^{+} \to K^{+} \mu^{\pm} e^{\mp}$ decays has been performed
- ⇛ Over a order of magnitude improvement wrt. to BaBar results
- ⇛ We target PRL.
- ⇛ No significant excess of events has been observed
- ⇛ Proponents ask the collaboration to approve this analysis.

Many thanks to the RC for their hard work and comments!
Many thanks to Stephane for fast reading the draft.

Backup

Combined banana



So large errors

⇛ $BR = 4.2 \times 10^{-9}$ and $BR = 4.5 \times 10^{-9}$.