HEPLike - tool for experimental likelihood evaluation

Jihyun Bhom
jihyun.bohm@cern.ch
Marcin Chrzaszcz
mchrzasz@cern.ch

(Re)interpreting the results of new physics searches at the LHC,
London, April 2-4, 2019

HEP results

⇛ How do we publish results?

CERN-EP-2019-043
LHCb-PAPER-2019-009
22 March 2019

Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays

LHCb collaboration†

Abstract

A measurement of the ratio of branching fractions of the decays $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ is presented. The proton-proton collision data used correspond to an integrated luminosity of $5.0\,\mathrm{fb}^{-1}$ recorded with the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV. For the dilepton mass-squared range $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ the ratio of branching fractions is measured to be $R_K = 0.846\,^{+0.060}_{-0.054}\,^{+0.016}_{-0.014}$, where the first uncertainty is statistical and the second systematic. This is the most precise measurement of R_K to date and is compatible with the Standard Model at the level of 2.5 standard deviations.

Submitted to Phys. Rev. Lett.

†Authors are listed at the end of this paper.

HEP results

⇛ How do we publish results?

CERN-EP-2019-043
LHCb-PAPER-2019-009
22 March 2019

arXiv:1903.09252v1 [hep-ex] 21 Mar 2019

Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays

LHCb collaboration†

Abstract

A measurement of the ratio of branching fractions of the decays $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ is presented. The proton-proton collision data used correspond to an integrated luminosity of $5.0\,\mathrm{fb}^{-1}$ recorded with the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV. For the dilepton mass-squared range $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ the ratio of branching fractions is measured to be $R_K = 0.846\,^{+0.060}_{-0.054}\,^{+0.016}_{-0.014}$, where the first uncertainty is statistical and the second systematic. This is the most precise measurement of R_K to date and is compatible with the Standard Model at the level of 2.5 standard deviations.

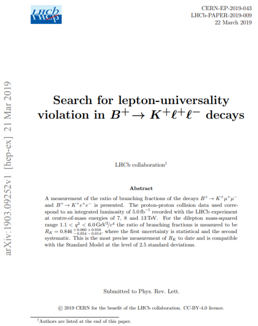
Submitted to Phys. Rev. Lett.

†Authors are listed at the end of this paper.

$$R_K = 0.846^{+0.060+0.016}_{-0.054-0.014}$$

HEP results

⇛ How do we publish results?


CERN-EP-2019-043
LHCb-PAPER-2019-009
22 March 2019

arXiv:1903.09252v1 [hep-ex] 21 Mar 2019

Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays

LHCb collaboration†

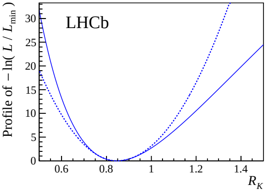
Abstract

A measurement of the ratio of branching fractions of the decays $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ is presented. The proton-proton collision data used correspond to an integrated luminosity of 5.0 fb^{-1} recorded with the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV. For the dilepton mass-squared range $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ the ratio of branching fractions is measured to be $R_K = 0.846\,^{+0.060}_{-0.054}\,^{+0.016}_{-0.014}$, where the first uncertainty is statistical and the second systematic. This is the most precise measurement of R_K to date and is compatible with the Standard Model at the level of 2.5 standard deviations.

Submitted to Phys. Rev. Lett.

© 2019 CERN for the benefit of the LHCb collaboration. CC-BY-4.0 licence.

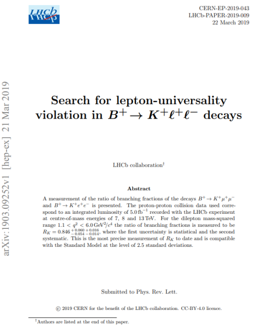
†Authors are listed at the end of this paper.



$$R_K = 0.846^{+0.060+0.016}_{-0.054-0.014}$$

HEP results

⇛ How do we publish results?


arXiv:1903.09252v1 [hep-ex] 21 Mar 2019

CERN-EP-2019-043
LHCb-PAPER-2019-009
22 March 2019

Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays

LHCb collaboration[†]

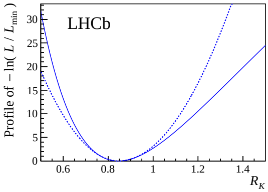
Abstract

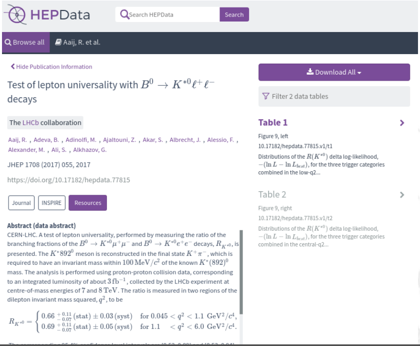
A measurement of the ratio of branching fractions of the decays $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ is presented. The proton-proton collision data used correspond to an integrated luminosity of $5.0\,\mathrm{fb}^{-1}$ recorded with the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV. For the dilepton mass-squared range $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ the ratio of branching fractions is measured to be $R_K = 0.846\,^{+0.060}_{-0.054}\,^{+0.016}_{-0.014}$, where the first uncertainty is statistical and the second systematic. This is the most precise measurement of R_K to date and is compatible with the Standard Model at the level of 2.5 standard deviations.

Submitted to Phys. Rev. Lett.

© 2019 CERN for the benefit of the LHCb collaboration. CC-BY-4.0 licence.

[†]Authors are listed at the end of this paper.




HEPData
Search HEPData
Search
Browse all
Aaij, R. et al.

Hide Publication Information

Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays

The LHCb collaboration

Aaij, R. , Adeva, B. , Adinolfi, M. , Ajaltouni, Z. , Akar, S. , Albrecht, J. , Alessio, F. , Alexander, M. , Ali, S. , Alkhazov, G.

JHEP 1708 (2017) 055, 2017

https://doi.org/10.17182/hepdata.77815

Journal | INSPIRE | Resources

Abstract (data abstract)

CERN-LHC. A test of lepton universality, performed by measuring the ratio of the branching fractions of the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, $R_{K^{*0}}$, is presented. The $K^{*}892^0$ meson is reconstructed in the final state $K^+\pi^-$, which is required to have an invariant mass within $100\,\mathrm{MeV}/c^2$ of the known $K^*(892)^0$ mass. The analysis is performed using proton-proton collision data, corresponding to an integrated luminosity of about $3\,\mathrm{fb}^{-1}$, collected by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The ratio is measured in two regions of the dilepton invariant mass squared, q^2, to be

$$R_{K^{*0}} = \begin{cases} 0.66\,^{+0.11}_{-0.07}\,(\mathrm{stat}) \pm 0.03\,(\mathrm{syst}) & \text{for } 0.045 < q^2 < 1.1\ \mathrm{GeV}^2/c^4, \\ 0.69\,^{+0.11}_{-0.07}\,(\mathrm{stat}) \pm 0.05\,(\mathrm{syst}) & \text{for } 1.1 < q^2 < 6.0\ \mathrm{GeV}^2/c^4. \end{cases}$$

Download All

Filter 2 data tables

Table 1
Figure 9, left
10.17182/hepdata.77815.v1/t1
Distributions of the $R(K^{*0})$ delta log-likelihood, $-(\ln L - \ln L_{best})$, for the three trigger categories combined in the low-q2...

Table 2
Figure 9, right
10.17182/hepdata.77815.v1/t2
Distributions of the $R(K^{*0})$ delta log-likelihood, $-(\ln L - \ln L_{best})$, for the three trigger categories combined in the central-q2...

$$R_K = 0.846^{+0.060+0.016}_{-0.054-0.014}$$

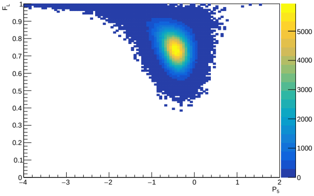
HEP results

⇛ How are the results used?

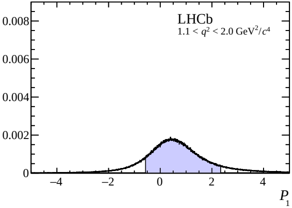
⇛ Correlations are neglected

	F_L	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_L	1.00	−0.13	−0.14	0.01	−0.03	0.10	−0.03	−0.01
S_3		1.00	−0.06	0.09	0.07	−0.02	0.01	−0.07
S_4			1.00	−0.19	−0.09	−0.05	0.12	0.07
S_5				1.00	−0.01	0.05	−0.02	0.10
A_{FB}					1.00	−0.01	−0.10	0.10
S_7						1.00	0.07	−0.05
S_8							1.00	−0.01
S_9								1.00

⇛ Non Linear effects are forgotten



⇛ Errors are being symmetrized



HEP results

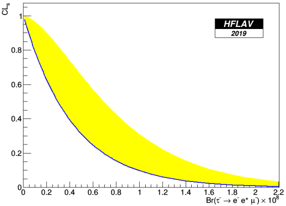
⇛ How are the results used?

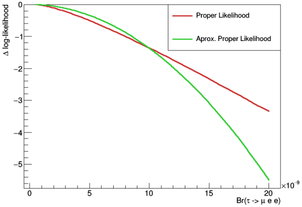
⇛ Interpreting Upper limits [HLFAV, 90% UL]:

$$\mathcal{B}(\tau \to \mu\mu e) < 9.9 \times 10^{-9}$$

⇛ People interpret this assuming it's a gaussian centered around 0 and width $\frac{9.9\times10^{-9}}{1.64}$.

⇛ Usually a full p-value scan is published:





⇛ The examples go on and on...

The idea

⇛ The theory and experimental community need to work together about proper interpretation.

The idea

$\Rightarrow$ The theory and experimental community need to work together about proper interpretation.

HEPLike

⇛ High Energy Physics Likelihood (HEPLike).

- Open source software.
- With separate database of measurements.
- Statistics library.
- Can be interfaced with existing codes.

⇛ It constructs the experimental likelihoods for you!
⇛ Does work with both the χ^2 and (log-)likelihood fits.
⇛ Useful utilities for creating citations and database search.

HEPLike

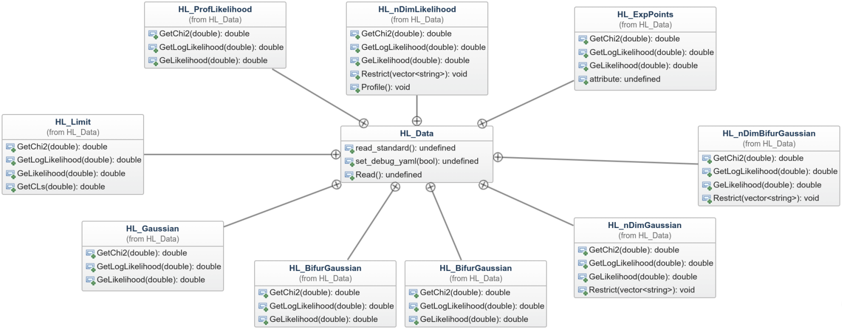
⇛ The are couple of measurement types:

- Upper limits,
- Single measurement with symmetric uncertainty,
- Single measurement with asymmetric uncertainty,
- Multiple measurements with symmetric uncertainty,
- Multiple measurements with asymmetric uncertainty,
- One dimensional likelihood function,
- n-dimensional likelihood function.

Bonus

In addition we provide a way for the future that the experiments can publish the dataset.

HEPLike - code structure



Measurement encoding, `Hl_Data`

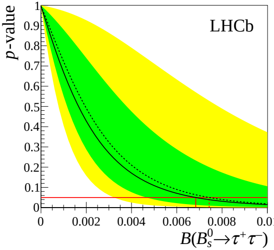
⇛ Measurements are stored in `YAML` file:


```
BibCite: Aaij:2017vbb
BibEntry: '@article{Aaij:2017vbb,
      author         = "Aaij, R. and others",
      title          = "{Test of lepton universality
                        with $B^{0} \rightarrow
                       K^{*0}\ell^{+}\ell^{-}$ decays}",
      collaboration  = "LHCb",

      }
      '
 DOI: 10.1007/JHEP08(2017)055
 Process: R_{Kstar^{*}}
 FileName: RKstar.yaml
 Name: RKstar
 Source: HEPDATA
 SubmissionYear: 2017
 PublicationYear: 2018
 Arxiv: 1705.05802
 Collaborations: LHCb
 Kinematics: q2>1.1 && q2<6.
 HLAuthor: Gal Anonim
 HLEmail: gal.anonim@ifj.edu.pl
 HLType: HL_ProfLikelihood
```

Upper limits, `HL_Limit`

⇛ Example of published p-value scans:



⇛ Information coded as:

```
Cls:
- [0.0, 1.0]
- [1.0e-10, 0.977091694706]
- [2.0e-10, 0.954375824297]
- [3.0e-10, 0.93200355343]
- [4.0e-10, 0.910630700546]
- [5.0e-10, 0.889382721809]
```

Upper limits, `HL_Limit`

$$pdf(x) = \frac{1}{2^{1/2}\Gamma(1/2)} x^{1/2-1} e^{-x/2}, \tag{1}$$

which had the cumulative distribution function defined as:

$$cdf(x) = \frac{1}{\Gamma(1/2)} \gamma(1/2, x/2). \tag{2}$$

In the above equations the $\Gamma(x)$ and $\gamma(k, x)$ correspond to Gamma and incomplete gamma functions. By revering the $cdf(x)$ one can obtain the χ^2 value:

$$\chi^2 = cdf^{-1}(1-p), \tag{3}$$

and if needed the log-likelihood:

$$-\log(\mathcal{L}) = \frac{1}{2}\chi^2, \tag{4}$$

Single measurement, symmetric error, `HL_Gaussian`

⇛ Well this is as simple as:

```
Observables:
- [ "Br_A2BCZ", 0.1, 0.05, 0.01 ]
```

⇛ The χ^2 is simple:

$$\chi^2 = \frac{(x_{obs} - x)^2}{\sigma_{stat}^2 + \sigma_{syst}^2}, \tag{5}$$

⇛ Wilks theorem can be used to translate to (log-)likelihood.

Single measurement, symmetric error, `HL_Gaussian`

⇛ Well this is as simple as:

```
Observables:
- [ "Br_A2BCZ", 0.1, 0.05, 0.01 ]
```

⇛ The χ^2 is simple:

$$\chi^2 = \frac{(x_{obs} - x)^2}{\sigma_{stat}^2 + \sigma_{syst}^2}, \tag{6}$$

⇛ Wilks theorem can be used to translate to (log-)likelihood.

Multiple measurement, symmetric error, `HL_nDimGaussian`

⇛ You need to pass two arguments:

```
Observables:
- [ "BR1", 0.1, 0.02]
- [ "BR2", 0.2, 0.01, 0.01]
- [ "BR3", 0.4, 0.04]
Correlation:
-   [ "BR1", "BR2", "BR3"]
-   [ 1.  ,  0.2 ,  0     ]
-   [ 0.2,   1.,    0.  ]
-   [ 0  ,   0.,   1.   ]
```

⇛ From this one constructs the covariance matrix, and evaluates the χ^2:

$$\chi^2 = V^T \mathrm{Cov}^{-1} V, \tag{7}$$

Measurement, asymmetric error, `HL_BifurGaussian`, `HL_ndimBifurGaussian`

⇛ You need to pass two arguments:

```
Observables:
- [ "BR1", 0.1, +0.02, -0.01, 0.02]
- [ "BR2", 0.2, +0.01, -0.05, +0.03, -0.02]
- [ "BR3", 0.3, +0.04, -0.03, 0.05]
Correlation:
-  [ "BR1", "BR2", "BR3"]
-  [ 1. ,   0.1 ,   0.2    ]
-  [ 0.1,    1.,    0.1 ]
-  [ 0.2 ,    0.1,   1.   ]
```

⇛ We choose to interpret this as Bifurcated Gaussian:

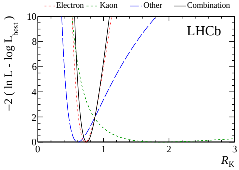
$$\mathrm{Cov}_{i,j} = \begin{cases} \mathrm{Corr}_{i,j}\ \sigma^i_+\sigma^j_+, & \text{if } x^i \geqslant x^i_{obs} \text{ and } x^j \geqslant x^j_{obs} \\ \mathrm{Corr}_{i,j}\ \sigma^i_+\sigma^j_-, & \text{if } x^i \geqslant x^i_{obs} \text{ and } x^j < x^j_{obs} \\ \mathrm{Corr}_{i,j}\ \sigma^i_-\sigma^j_+, & \text{if } x^i < x^i_{obs} \text{ and } x^j \geqslant x^j_{obs} \\ \mathrm{Corr}_{i,j}\ \sigma^i_-\sigma^j_-, & \text{if } x^i < x^i_{obs} \text{ and } x^j < x^j_{obs} \end{cases} \quad (8)$$

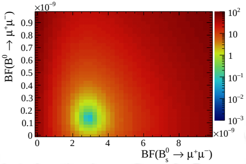
Likelihoods, `HL_ProfLikelihood`, `HL_nDimLikelihood`

⇛ Here we add just the location of `ROOT` object.

```
ROOTData: data/HEPData-ins1599846-v1-Table_1.root
TGraphPath: "Table 1/Graph1D_y1"
```

```
ROOTData: data/LHCb/RD/Bs2mumu_5fb/histB2mumu.root
TH2Path: "h_2DScan"
```

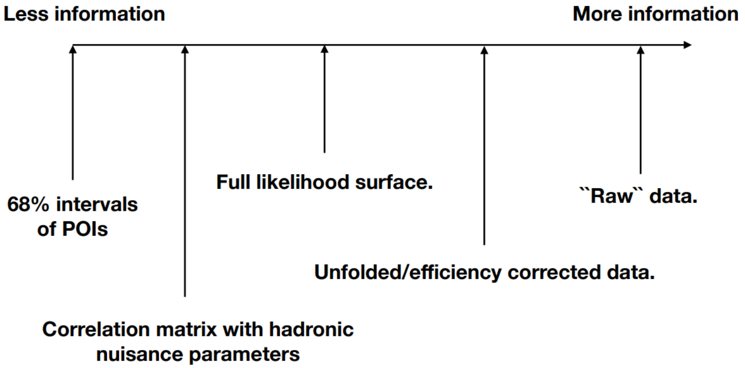




⇛ This is the best way to publish results!!!

⇛ The problem is in what way one should publish the higher dim likelihoods?

Publishing data



Stolen from P. Owen

Publishing data HL_ExpData

⇛ The YAML entry:

```
ROOTData: data/toy/data.root
TTreePath: t
Observables:
- [ x ]
- [ y ]
- [ z ]
Weight: w
```

⇛ Set the PDF you want to fit:

```
double (*fun)(vector<double> par , vector<double> point)
```

⇛ The program will evaluate the (log-)likelihood on the whole dataset for given parameters.

⇛ You only need a scanning tools and you are done.

Useful functions

⇛ Search for measurement you need:

```
python lookup.py --Arxiv 1705.05802
Found files:
../data/examples/RKstar_lowq2.yaml
```

⇛ Create citation file:

```
Aaij:2017vbb
b2mumu.yaml
```

To prepare the `BiBtex` file user should run the `make_citations.py` script located in the `utils` directory:

```
cd utils
python make_citations.py list.txt
```

Check it out

⇛ The HEPLike code:
`https://github.com/mchrzasz/HEPLike`

⇛ The HEPLike database:
`https://github.com/mchrzasz/HEPLikeData`

Don't be shy! Give it a spin. Feedback is welcomed.

Check it out

⇛ The HEPLike code:
`https://github.com/mchrzasz/HEPLike`
⇛ The HEPLike database:
`https://github.com/mchrzasz/HEPLikeData`

Don't be shy! Give it a spin. Feedback is welcomed.

Thank you for your attention

Backup