$R(K^*)$ measurement in ATLAS

NYUAD and WIS Collaboration

22. 12. 2020

Tomáš Jakoubek on behalf of the analysis team

Weizmann Institute of Sciences | tomas.jakoubek@cern.ch

Introduction

- Searching for deviations from the Standard Model (SM) expectations: **Lepton Flavour Universality** (LFU) - the couplings of the charged leptons to the gauge bosons are equal.
- Using rare decays $B_d^0 \to K^* e^+ e^-$ and $B_d^0 \to K^* \mu^+ \mu^-$.

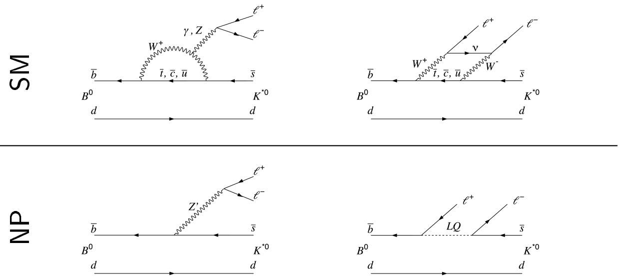


Figure: Taken from [1].

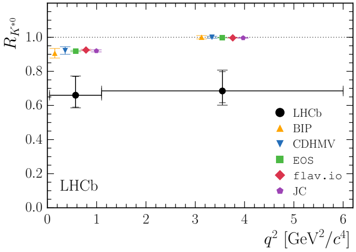


Figure: Taken from [1].

- LHCb result from Run1 (3 fb^{-1}) [1]: compatible with the SM expectations at the level of $2.4 - 2.5\ \sigma$.
- Tensions can be seen, especially if combined with $R(K)$ measurement ($3 - 4\ \sigma$) [2].

Experiment

- In ratios, hadronic uncertainties of the theoretical predictions cancel. The ratio

$$R(K^*) = \left.\frac{\mathcal{B}(B_d^0 \to K^* \mu\mu)}{\mathcal{B}(B_d^0 \to K^* ee)}\right|_{q^2}$$

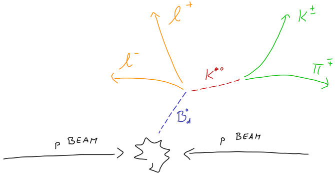
deviates from 1 only because of m_e and m_μ.

- Experiments: measuring double ratio reduces systematic uncertainties

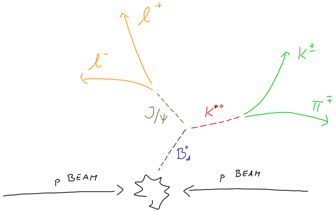
$$R(K^*) = \frac{\mathcal{B}(B_d^0 \to K^* \mu\mu)}{\mathcal{B}(B_d^0 \to K^* J/\psi(\to \mu\mu))} \cdot \left.\frac{\mathcal{B}(B_d^0 \to K^* J/\psi(\to ee))}{\mathcal{B}(B_d^0 \to K^* ee)}\right|_{q^2},$$

i.e., measure yields and efficiencies for the resonant and non-resonant modes.

- **In ATLAS:** completely driven by *ee*-part (both *analysis*- and *trigger*-side).



vs.



$R(K^*)$ measurement @ ATLAS
Electron triggers

- For $R(K^*)$, we need to trigger low p_T di-electron events.
- New triggers developed during the later stage of Run2 data-taking, recorded $\sim 40\ \text{fb}^{-1}$ of *pp* data since the deployment.
  - Measurement would not be possible without them!
  - **Unseeded** chains $\rightarrow$ running on every Level-1 accepted event!
- Proper evaluation and efficiency study ongoing (needed also for the future trigger development).

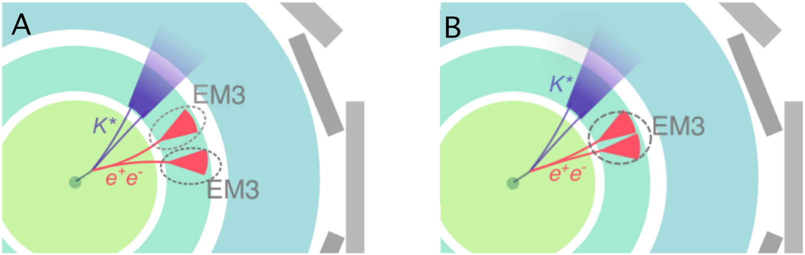


Figure: H. Russell

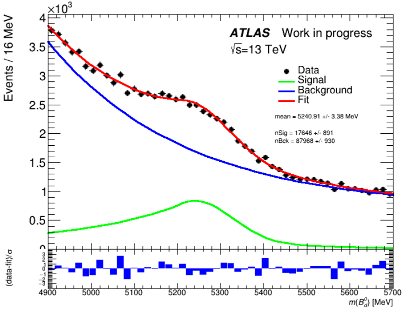
Data 2018

Non-resonant estimate from resonant data

- To extrapolate from the resonant to non-resonant channel (currently **blinded**):

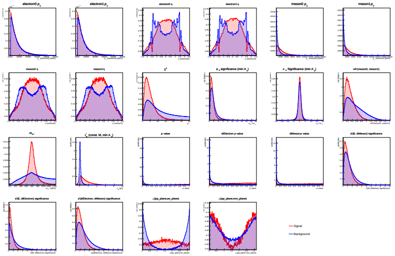
$$n^{\text{obs}}_{\text{Non}} = n^{\text{obs}}_{\text{Res}} \cdot \frac{\mathcal{B}_{\text{Non}} \cdot [A \times \epsilon]_{\text{Non}}}{\mathcal{B}_{\text{Res}} \cdot [A \times \epsilon]_{\text{Res}}}.$$

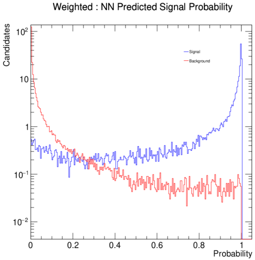
- The estimate for the "LHCb bin" based on the *preliminary* **cut-based** selection in the resonant channel is $n^{\text{obs}}_{\text{Non}} \approx 100$.
- LHCb has $n^{\text{obs}}_{\text{Non}} = 188 \pm 27$ in approximately the same range [1].
- Not possible to use the same procedure in the non-resonant channel.
- Instead of cut-based selection, using **machine learning** to improve the signal selection efficiency.

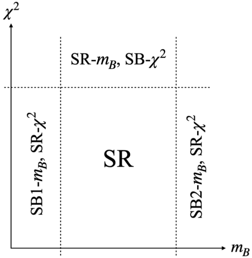


Machine learning strategy

- Using Neural Network (NN) with features including momenta, masses, vertexing information, angles etc.
- Data-driven background.
- Training and testing in different regions.
- Trying two NNs (for targeting the combinatorial background and for targeting the real peaking processes.
- We expect about a factor of 2 improvement in the signal efficiency (optimization of NN and inclusion of new features are in development).



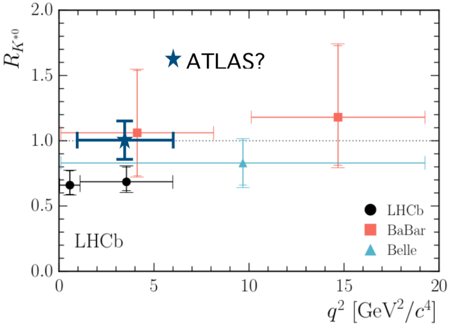




Summary

- Important to join the LFU efforts with the **first** measurement of $R(K^*)$ at **ATLAS**. Both LHCb and CMS are working on Run-2 analyses.
- Very challenging, since ATLAS is a general purpose detector, not dedicated to B-physics.
- Electron trigger developments in 2018 enable us to do this analysis in Run-2.
- Analysis **unique** in many ways!
- Lead by the **WIS team**.
- Machine learning approach is very promising.
- Estimate of the statistical uncertainty in the measurement of $R(K^*)$: competitive to LHCb!

THANK YOU!



R(K*) plotted at 1 due to SM expectations

BACKUP

$R(K^*)$ measurement @ ATLAS
Electron triggers

- For $R(K^*)$, we need to trigger low p_{T} di-electron events.
- New triggers developed during the later stage of Run2 data-taking.
  - Measurement would not be possible without them!
  - e-seeded, μ-seeded, and **unseeded** chains $\rightarrow$ running on every Level-1 (L1) accepted event!
- 3 GeV L1 EM regions of interest, two 5 GeV electrons with very loose ID on the higher trigger level.
- Di-electron vertex with 0.1 GeV $< m_{ee} <$ 6 GeV.
- Deployed on 14 July 2018, taken $\sim$ 40 fb^{-1} of pp data.
- Proper evaluation and efficiency study ongoing (needed also for the future trigger development).

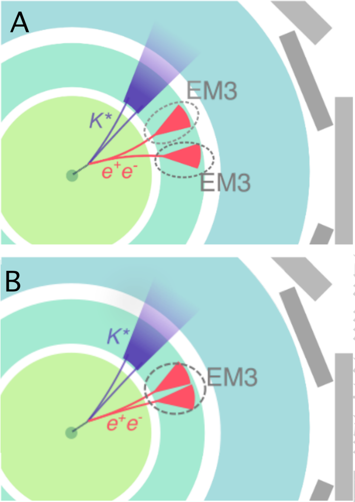
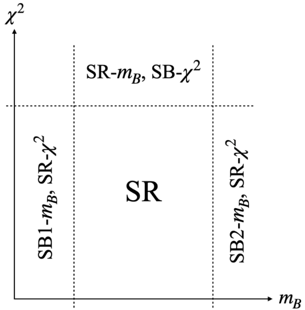


Figure: H. Russell

Machine learning

- Using Neural Network (NN) with features including momenta, masses, vertexing information, angles etc.
- Trying two NNs:
  - NN1: To target combinatorial background (using signal MC in B_d^0 mass signal-region and real data in B_d^0 mass sidebands).
  - NN2: To target real peaking processes (using NN1 filtered MC/real data in B_d^0 mass signal-region in χ^2 signal-region/sideband).
- ML algorithm specifications:
  - Classifer type: Neural network,
  - Optimizer: Adam,
  - Loss: Binary crossentropy,
  - Evaluation metrics: Accuracy,
  - Activation function: Sigmoid.



NN performance (WIP)

And some implications

	Signal events [%]	Background events [%]
mB_SR+chi2_SR	94.76	73.85
After NN1 @ 0.2	86.49	16.54
After NN2 @ 0.1	84.44	11.75
After NN2 @ 0.2	80.91	8.86
After NN2 @ 0.3	76.52	6.81
After NN2 @ 0.4	70.68	5.11
After NN2 @ 0.5	63.56	3.59
After NN2 @ 0.6	55.13	2.28
After NN2 @ 0.7	44.48	1.14
After NN2 @ 0.8	31.73	0.40
After NN2 @ 0.9	18.97	0.13
Full cut-based selection	43.71	2.50

- From NN performance, we can expect about a factor of 2 improvement in signal efficiency with a similar background rejection (optimization of NN and inclusion of new features are in development).
- We estimate the statistical uncertainty in the measurement of $R(K^*)$ to be $< 15\%$.
- For LHCb, the statistical uncertainty is 10 % – 16 %.

Partial selection

- for MC: test isTrue
- q(e0)*q(e1) < 0 and q(m0)*q(m1) < 0
- dR(e0,e1) > 0.1
- pT(e{0,1}) > 5 GeV
- eta(e{0,1}) < 2.5
- pT(m{0,1}) > 500 MeV
- eta(m{0,1}) < 2.5
- m(ee) < 7 GeV
- 3 GeV < m(B) 6.5 GeV or 3 GeV < m(BBar) 6.5 GeV
- 690 MeV < m(piK) < 1110 MeV or 690 MeV < m(Kpi) < 1110 MeV

R(K*) Analysis in CMS

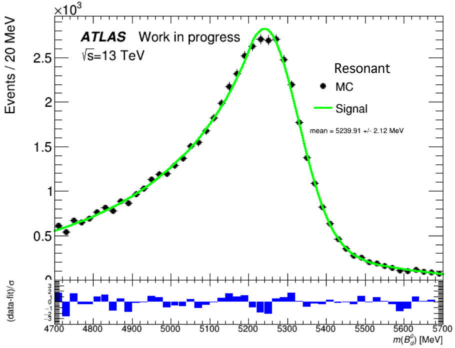
- CMS estimates to have 2600 $B^0 \to K^* \ell\ell$ events, before any fiducial cuts [3].

Our Guess for the Observed Number of Events at CMS

- Using $1\ \mathrm{GeV}$ cuts for electrons, and tracks (based on their reconstruction ability), and a generous η cut at 2.7, we can estimate a filter efficiency of $\sim 24\,\%$
- Taking into account efficiencies, an average of ~ 0.7 for electrons [3], and generously assuming the same efficiency for both tracks, the observed number of events $\approx 2600 \cdot 0.24 \cdot 0.7^4 = 150$.
- Further, assuming a high percentage of events passing vertexing, identification, etc. ($\sim 90\,\%$), and assuming a high selection efficiency ($\sim 80\,\%$), we can estimate the observed number of events to be < 110.

Cut-based Selection

- Electron channel is much more challenging, especially the non-resonant channel.
- Very loose selection on the derivation level:
  - $2e$ vertexing and $2e + 2\mathrm{trks}$ vertexing, di-track mass cut around m_{K*}.
- Current (preliminary) cuts:
  - $p_T(e) > 5$ GeV, $p_T(\mathrm{trk}) > 1$ GeV
  - $|\eta| < 2.5$
  - Four-track vertex $\chi^2/\mathrm{nDoF} < 2$
  - $\tau(B_d) > 0.2$ ps
  - $|m(K\pi) - m(K^*)| < 50$ MeV
  - 4700 MeV $< m(B_d) < 5700$ MeV
  - Selecting the best χ^2/nDoF candidate in the event.
- Fit the m_B spectrum in the resonant and non-resonant channels.
  - Unbinned maximum likelihood fit.
  - Sum of Crystal-Ball and Gaussian functions with a common mean, all parameters are free.



m_B spectrum fit for the resonant channel (MC)

Estimation of Observed Non-resonant Events from Resonant Data

Reproduced fron Noam's Talk in the R(K*) Meeting on 29 June 2020 [6]

In general:

(1) $$\mathscr{B}_{X\to Y} = \frac{n^{\text{tru}}_{X\to Y}}{n_{X\to\text{everything}}} = \frac{n^{\text{obs}}_{X\to Y}/[\mathscr{A}\times\epsilon]_{X\to Y}}{n_{X\to\text{everything}}}$$

(2) $$\frac{\mathscr{B}_{X\to Y}}{\mathscr{B}_{X\to Z}} = \frac{n^{\text{obs}}_{X\to Y}/[\mathscr{A}\times\epsilon]_{X\to Y}}{n^{\text{obs}}_{X\to Z}/[\mathscr{A}\times\epsilon]_{X\to Z}}$$ since $n_{X\to\text{everything}}$ cancels in the ratio.

(3) $$\frac{n^{\text{obs}}_{X\to Y}}{n^{\text{obs}}_{X\to Z}} = \frac{\mathscr{B}_{X\to Y}\cdot[\mathscr{A}\times\epsilon]_{X\to Y}}{\mathscr{B}_{X\to Z}\cdot[\mathscr{A}\times\epsilon]_{X\to Z}}$$

And so:

(4) $$\frac{n^{\text{obs}}_{\text{Non}}}{n^{\text{obs}}_{\text{Res}}} = \frac{\mathscr{B}_{\text{Non}}\cdot[\mathscr{A}\times\epsilon]_{\text{Non}}}{\mathscr{B}_{\text{Res}}\cdot[\mathscr{A}\times\epsilon]_{\text{Res}}}$$

(4) $\frac{n^{obs}_{Non}}{n^{obs}_{Res}} = \frac{\mathscr{B}_{Non}}{\mathscr{B}_{Res}} \cdot \frac{[\mathscr{A} \times \epsilon]_{Non}}{[\mathscr{A} \times \epsilon]_{Res}} \implies n^{obs}_{Non} = n^{obs}_{Res} \cdot \frac{\mathscr{B}_{Non}}{\mathscr{B}_{Res}} \cdot \frac{[\mathscr{A} \times \epsilon]_{Non}}{[\mathscr{A} \times \epsilon]_{Res}}$

- (4)(1) $\mathscr{B}_{Non} = \mathscr{B}(B \to K^* ee)$ from PDG
- (4)(2) $\mathscr{B}_{Res} = \mathscr{B}(B \to K^* J/\psi) \cdot \mathscr{B}(J/\psi \to ee)$ both from PDG
- (4)(3) $[\mathscr{A} \times \epsilon]_X = n^{sel}_X / n^{gen}_X$ in the full q^2 range!
  - (4)(3)(1) n^{sel}_X from running the full selection on X-signal Monte-Carlo
  - (4)(3)(2) $n^{gen}_X = n^{AOD}_X / \epsilon^{gen}_X$, where n^{AOD}_X and ϵ^{gen}_X are from AMI (note: not DAOD!)

(5) $n^{obs}_{Non} = n^{obs}_{Res} \cdot \frac{\mathscr{B}(B \to K^* ee)}{\mathscr{B}(B \to K^* J/\psi) \cdot \mathscr{B}(J/\psi \to ee)} \cdot \frac{n^{sel}_{Non} / (n^{AOD}_{Non} / \epsilon^{gen}_{Non})}{n^{sel}_{Res} / (n^{AOD}_{Res} / \epsilon^{gen}_{Res})}$

- (5)(1) all quantities here correspond to the full range of q^2 and this range must be identical between the two channels (Non and Res)!
- (5)(2) n^{obs}_{Res} is obtained from a fit of the Res events (in the full range of q^2) in data to the $s + b$ model in the m_B distribution and integrating over the s component

(6) $n^{obs}_{Non}(q^2 < q^2_0\ \text{GeV}^2) = n^{obs}_{Non} \cdot f_{Non}(q^2 < q^2_0\ \text{GeV}^2)$

- (6)(1) n^{obs}_{Non} is given by Eq.(5) in the full q^2 range!
- (6)(2) $f_{Non}(q^2 < 6\ \text{GeV}^2) \sim 0.347$ [TBC] is the truth-level fraction of events found below 6 GeV^2. Resolution leads to bin-migration but this works equally in both directions.

Rearranging for $B + \bar{B}$:

Starting from equations (4)(3) and (4)(3)(1) + (4)(3)(2) we need to correct for the contributions form the two conjugated processes

$$(6)\quad n_{\mathrm{X}}^{\mathrm{gen}} = n_{\mathrm{X},B}^{\mathrm{gen}} + n_{\mathrm{X},\bar{B}}^{\mathrm{gen}} = n_{\mathrm{X},B}^{\mathrm{AOD}}/\epsilon_{\mathrm{X},B}^{\mathrm{gen}} + n_{\mathrm{X},\bar{B}}^{\mathrm{AOD}}/\epsilon_{\mathrm{X},\bar{B}}^{\mathrm{gen}} = \frac{n_{\mathrm{X},B}^{\mathrm{AOD}}\epsilon_{\mathrm{X},\bar{B}}^{\mathrm{gen}} + n_{\mathrm{X},\bar{B}}^{\mathrm{AOD}}\epsilon_{\mathrm{X},B}^{\mathrm{gen}}}{\epsilon_{\mathrm{X},B}^{\mathrm{gen}}\epsilon_{\mathrm{X},\bar{B}}^{\mathrm{gen}}}$$

$$(7)\quad n_{\mathrm{X}}^{\mathrm{sel}} = n_{\mathrm{X},B}^{\mathrm{sel}} + n_{\mathrm{X},\bar{B}}^{\mathrm{sel}}$$

Finally, putting it all together:

$$(8)\ n_{\mathrm{Non}}^{\mathrm{obs}}(q_0^2) = n_{\mathrm{Res}}^{\mathrm{obs}} \cdot \frac{\mathscr{B}(B \to K^*ee)}{\mathscr{B}(B \to K^*J/\psi)\cdot\mathscr{B}(J/\psi \to ee)} \cdot \frac{(n_{\mathrm{Non},B}^{\mathrm{sel}} + n_{\mathrm{Non},\bar{B}}^{\mathrm{sel}}) \cdot \frac{\epsilon_{\mathrm{Non},B}^{\mathrm{gen}}\epsilon_{\mathrm{Non},\bar{B}}^{\mathrm{gen}}}{n_{\mathrm{Non},B}^{\mathrm{AOD}}\epsilon_{\mathrm{Non},\bar{B}}^{\mathrm{gen}} + n_{\mathrm{Non},\bar{B}}^{\mathrm{AOD}}\epsilon_{\mathrm{Non},B}^{\mathrm{gen}}}}{(n_{\mathrm{Res},B}^{\mathrm{sel}} + n_{\mathrm{Res},\bar{B}}^{\mathrm{sel}}) \cdot \frac{\epsilon_{\mathrm{Res},B}^{\mathrm{gen}}\epsilon_{\mathrm{Res},\bar{B}}^{\mathrm{gen}}}{n_{\mathrm{Res},B}^{\mathrm{AOD}}\epsilon_{\mathrm{Res},\bar{B}}^{\mathrm{gen}} + n_{\mathrm{Res},\bar{B}}^{\mathrm{AOD}}\epsilon_{\mathrm{Res},B}^{\mathrm{gen}}}} \cdot f_{\mathrm{Non}}(q_0^2)$$

Notes:

- the fits to the m_B distribution to get $n_{\mathrm{Non},B}^{\mathrm{sel}}$, $n_{\mathrm{Non},\bar{B}}^{\mathrm{sel}}$, $n_{\mathrm{Res},B}^{\mathrm{sel}}$, and $n_{\mathrm{Res},\bar{B}}^{\mathrm{sel}}$ can be done separately on the Monte-Carlo signals, but in fact, there's no need to do the fits on MC and it is enough to count events
- the fit to get $n_{\mathrm{Res}}^{\mathrm{obs}}$ in data cannot be done separately for B and $\bar{B}$, surely not with the cut-based option
- any ML-based improvement in the Res selection will be evident for both $n_{\mathrm{Res}}^{\mathrm{obs}}$ and $n_{\mathrm{Res},B}^{\mathrm{sel}} + n_{\mathrm{Res},\bar{B}}^{\mathrm{sel}}$ so it will cancel in the ratio. Therefore, we can just estimate the Res channel with the cut-based selection
- the ML improvement in $n_{\mathrm{Non},B}^{\mathrm{sel}} + n_{\mathrm{Non},\bar{B}}^{\mathrm{sel}}$ will not cancel. We can estimate this number with the cut-based selection and scale it up to the ML-based level or simply count the events passing this selection

Cut-based vs ML-based:

(9) $n^{\rm sel}_{{\rm Non},B} + n^{\rm sel}_{{\rm Non},\bar{B}}$ need to compare this number between the the cut-based and ML-based options.

(10) $n^{\rm sel\ (test-sample)}_{{\rm Non},Q} = \left((\sim 10\%) \times n^{\rm AOD}_{{\rm Non},Q} \times \epsilon^{\rm partial}_{{\rm Non},Q} \right) \times \epsilon^{\rm rest}_{{\rm Non},Q}$ for $Q = \{B, \bar{B}\}$ working only with the test sample of the NN, i.e. only ~10% of the sample, but really the same test-sample!

(10)(1) $n^{\rm sel\ (test-sample)}_{{\rm Non},Q}({\rm Cut-based}) = \left((\sim 10\%) \times n^{\rm AOD}_{{\rm Non},Q} \times \epsilon^{\rm partial}_{{\rm Non},Q} \right) \times \epsilon^{\rm cutbased}_{{\rm Non},Q}$

(10)(2) $n^{\rm sel\ (test-sample)}_{{\rm Non},Q}({\rm ML-based}) = \left((\sim 10\%) \times n^{\rm AOD}_{{\rm Non},Q} \times \epsilon^{\rm partial}_{{\rm Non},Q} \right) \times \epsilon^{\rm NN}_{{\rm Non},Q}$

(11) $\dfrac{n^{\rm sel\ (test-sample)}_{{\rm Non},Q}({\rm ML-based})}{n^{\rm sel\ (test-sample)}_{{\rm Non},Q}({\rm Cut-based})} = \dfrac{\epsilon^{\rm NN}_{{\rm Non},Q}}{\epsilon^{\rm cutbased}_{{\rm Non},Q}}$ this must hold also for the full sample so:

(11)(1) $\dfrac{n^{\rm sel}_{{\rm Non},Q}({\rm ML-based})}{n^{\rm sel}_{{\rm Non},Q}({\rm Cut-based})} = \dfrac{\epsilon^{\rm NN}_{{\rm Non},Q}}{\epsilon^{\rm cutbased}_{{\rm Non},Q}}$ and hence

(11)(1)(1) $n^{\rm sel}_{{\rm Non},Q}({\rm ML-based}) = n^{\rm sel}_{{\rm Non},Q}({\rm Cut-based}) \cdot \dfrac{\epsilon^{\rm NN}_{{\rm Non},Q}}{\epsilon^{\rm cutbased}_{{\rm Non},Q}}$

where simply $\epsilon^{\rm cutbased}_{{\rm Non},Q} = \dfrac{n^{\rm cutbased}_{{\rm Non},Q}}{n^{\rm partial}_{{\rm Non},Q}}$ and $\epsilon^{\rm NN}_{{\rm Non},Q} = \dfrac{n^{\rm NN}_{{\rm Non},Q}}{n^{\rm partial}_{{\rm Non},Q}}$, i.e. both measured on the MC with respect to the partial selection applied to the test sample.

(12) $\left(n^{\rm sel}_{{\rm Non},B} + n^{\rm sel}_{{\rm Non},\bar{B}} \right) |_{\rm ML-based} = n^{\rm sel}_{{\rm Non},B}({\rm Cut-based}) \cdot \dfrac{\epsilon^{\rm NN}_{{\rm Non},B}}{\epsilon^{\rm cutbased}_{{\rm Non},B}} + n^{\rm sel}_{{\rm Non},\bar{B}}({\rm Cut-based}) \cdot \dfrac{\epsilon^{\rm NN}_{{\rm Non},\bar{B}}}{\epsilon^{\rm cutbased}_{{\rm Non},\bar{B}}}$

(13) to complete the calculation, replace the $n^{\rm sel}_{{\rm Non},B} + n^{\rm sel}_{{\rm Non},\bar{B}}$ term in Eq.(8) with the term in Eq.(12)

Estimation of Uncertainty in R(K*)

- $\mathscr{B}_{\text{Res}} = \frac{n^{\text{obs}}_{\text{Res}}/[\mathscr{A}\times\epsilon]_{\text{Res}}}{n_{B\to\text{everything}}}$ and $\mathscr{B}_{\text{Non}} = \frac{n^{\text{obs}}_{\text{Non}}/[\mathscr{A}\times\epsilon]_{\text{Non}}}{n_{B\to\text{everything}}}$

- $R(K^*) = \frac{\mathscr{B}^{\mu\mu}_{\text{Non}}}{\mathscr{B}^{\mu\mu}_{\text{Res}}} \cdot \frac{\mathscr{B}^{ee}_{\text{Res}}}{\mathscr{B}^{ee}_{\text{Non}}} = \left[\frac{n^{\text{obs}}_{\text{Non}}/[\mathscr{A}\times\epsilon]_{\text{Non}}}{n^{\text{obs}}_{\text{Res}}/[\mathscr{A}\times\epsilon]_{\text{Res}}}\right]_{\mu\mu} \cdot \left[\frac{n^{\text{obs}}_{\text{Res}}/[\mathscr{A}\times\epsilon]_{\text{Res}}}{n^{\text{obs}}_{\text{Non}}/[\mathscr{A}\times\epsilon]_{\text{Non}}}\right]_{ee}$

- $R(K^*) = \frac{A_{\mu\mu}/B_{\mu\mu}}{C_{\mu\mu}/D_{\mu\mu}} \cdot \frac{C_{ee}/D_{ee}}{A_{ee}/B_{ee}} = \frac{A_{\mu\mu}D_{\mu\mu}C_{ee}B_{ee}}{C_{\mu\mu}B_{\mu\mu}A_{ee}D_{ee}}$

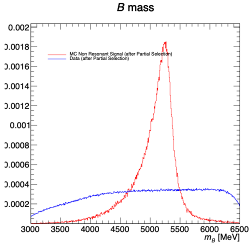
- $\frac{\Delta R(K^*)}{R(K^*)} = R(K^*) \cdot \sqrt{\left(\frac{\Delta A_{\mu\mu}}{A_{\mu\mu}}\right)^2 + \left(\frac{\Delta B_{\mu\mu}}{B_{\mu\mu}}\right)^2 + \left(\frac{\Delta C_{\mu\mu}}{C_{\mu\mu}}\right)^2 + \left(\frac{\Delta D_{\mu\mu}}{D_{\mu\mu}}\right)^2 + \left(\frac{\Delta A_{ee}}{A_{ee}}\right)^2 + \left(\frac{\Delta B_{ee}}{B_{ee}}\right)^2 + \left(\frac{\Delta C_{ee}}{C_{ee}}\right)^2 + \left(\frac{\Delta D_{ee}}{D_{ee}}\right)^2}$

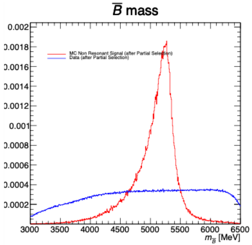
- Assuming $n^{\text{obs}}_{\text{Non},ee} = 100$, and $n^{\text{obs}}_{\text{Res},ee} \simeq 73 \times n^{\text{obs}}_{\text{Non},ee}$, and $n^{\text{obs}}_{\mu\mu} \simeq 2 \times n^{\text{obs}}_{ee}$ (at least) for both Res and Non, and $n^{\text{obs}}_{\text{Res},\mu\mu} \simeq 80 \times n^{\text{obs}}_{\text{Non},\mu\mu}$, and where 73 (80) is the ratio $\frac{\mathscr{B}_{\text{Non}}}{\mathscr{B}_{\text{Res}} \times \mathscr{B}_{J/\psi\to\ell\ell}}$ for electrons (muons).
- Assuming the relative uncerts on all other terms in the ratio are negligible
- Assuming that $R(K^*) \simeq 1$, we get $\frac{\Delta R(K^*)}{R(K^*)} = R(K^*) \cdot 12.2\,\% \simeq 12.2\,\%$
- To be conservative, we inflate this to 15 % .

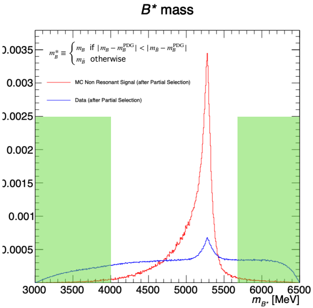
27

NN1: Sideband/Signal Region Defintions

- NN classification using the sideband data (2018) as background and MC non-resonant signal in signal region as signal.
- Define side-bands as $m_B < 4000$ MeV and $m_B > 5700$ MeV?
  - Four-track vertex reconstruction done twice, assuming $K\pi$ and πK. This means 2 reconstructed B-masses, m_B and $m_{\bar{B}}$.
  - Rather define the 'side-band' using a parameter which is symmetric under conjugation, m_B^*, and define the side-bands as the following ranges: $m_B* < 4000$ MeV and $m_B* > 5700$ MeV
  - Fraction of signal in side-bands is 3% (from MC).
- Notice that we are only using a small fraction of the data (~3%) here, and aren't using the signal-region data in any further analysis. So, bliding is not violated.

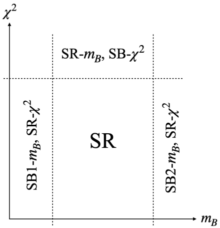


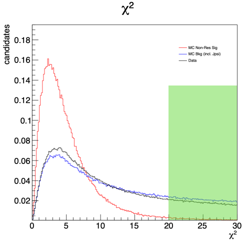




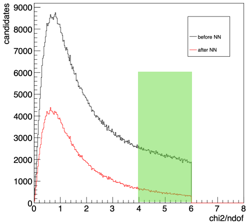
NN2: Sideband/Signal Region Definitions

- Looking at the $B-$mass signal region, we can identify a useful SB definition for data as $\chi^2 > 20$, for instance.
- Fraction of MC non-res signal in this SB: $1.7\,\%$
  - Fraction of MC non-res signal for $\chi^2 > 25 : 0.6\,\%$
  - Fraction of MC non-res signal for $\chi^2 > 15 : 4.2\,\%$
- The NN1 doesn't kill this SB scrupulously. So, this new classification wouldn't be redundant.
- We are only looking at a very tiny fraction of data, so looking at the χ^2 feature shouldn't be too much unblinding. Also, the NN1 cut is very loose (as seen in previous slide) so we are looking at the data very far from final selection.
- Notice that selecting a region in χ^2 doesn't limit us in the phase space of any other (kinematic) feature.





χ^2 : before and after NN1 cut



Both plots are in the SR of B-mass

30

REFERENCES

References

[1] LHCb Collaboration, "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays," JHEP **1708** (2017) 055.

[2] M. Algueró, B. Capdevila, A. Crivellin, S. Descotes-Genon, P. Masjuan, J. Matias, M. Novoa Brunet and J. Virto, "Emerging patterns of New Physics with and without Lepton Flavour Universal contributions," Eur. Phys. J. C **79** (2019) no.8, 714