

#### **ETH** zürich

# Search for Lepton Flavour Universality Violation at the CMS experiment in $B_c$ semileptonic decays

Federica Riti Zürich PhD Seminars 2022, 26 January 2023

#### **Lepton Flavors**

- The Standard Model (SM) of particle physics is the theory describing three of the four known fundamental forces of nature (except gravitational force), and the particles connected to them.
- In the SM there are three flavors of leptons:  $e, \mu, \tau$ , which differ only in mass.

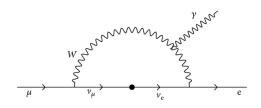


2 / 27

- To look for hints of new physics Beyond the Standard Model (BSM), violations to two of the leptons (accidental) SM symmetries can be searched:
  - Lepton Flavor Violation (LFV):
  - Lepton Flavor Universality Violation (LFUV)

### **Lepton Flavor Violation (LFV)**

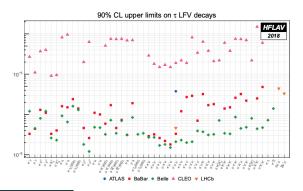
- The charged lepton flavor is conserved in the SM through an accidental symmetry.
- There is already evidence of neutral LFV, through neutrino oscillations
- Charged LFV just happens in loop diagrams with  $\nu$  mixing, and it is strongly suppressed in the SM (rate  $\sim 10^{-55}$ );
- BSM theories predict a larger Branching Ratio (BR) up to  $10^{-10}-10^{-8}$  EPJC57(2008)13-182



#### **LFV** - Few Experimental Examples

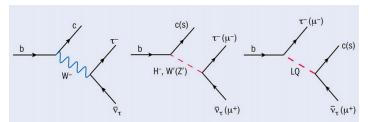
Direct LFV search can be done in different sectors:

- Higgs sector, with H o e au or  $H o \mu au$  decays PRD104(2021)032013
- ullet Top quark decays  $(t 
  ightarrow e \mu q)$  JHEP06(2022)082
- ullet Drell-Yann/ Z decays  $(Z 
  ightarrow (e/\mu au), Z(e \mu))$  arXiv:2204.10783
- B decays LHCb-TALK-2022-040
- Leptonic decays  $(\tau \rightarrow 3\mu)$  JHEP01(2021)163



#### **Lepton Flavor Universality Violation (LFUV)**

- Another accidental symmetry of the SM is that the three charged leptons have the same coupling constant in electro-weak (EW) interactions
- Therefore, if these couplings are found to be different, this would be a hint of physics BSM.
- What are the proposed explanations for a deviation from the SM?
- Existence of interactions involving new particles as:
  - Charged Higgs boson [1][2][3][4][5];
  - Leptoquarks (LQ) EXOtica;
  - New vector bosons (Summary plots EXO).



F.Riti

#### LFUV in the B Sector

- An interesting sector to look for LFUV is the B sector: decays involving hadrons made of b-quarks.
- The so-called R-measurements are amongst the most interesting that can be performed in this sector:

 $R(J/\psi)$  Analysis

$$R_{H_s} = rac{\mathcal{B}(H_b 
ightarrow H_s \mu^+ \mu^-)}{\mathcal{B}(H_b 
ightarrow H_s e^+ e^-)}$$

- $b \rightarrow sI^+I^-$
- loop level → smaller BR;
- $\nu$ -less  $\rightarrow$  fully reconstructed:
- precise predictions.



$$R_{H_c} = rac{\mathcal{B}(H_b 
ightarrow H_c au^- ar{
u_{ au}})}{\mathcal{B}(H_b 
ightarrow H_c \mu^- ar{
u_{ au}})}$$

- $b \rightarrow cl^-\bar{\nu}_l$
- tree level → large BR;
   sensitive to syst uncert;
- $\nu$ 's  $\rightarrow$  missing mass; template fits.
- sensitive to QCD calculations.

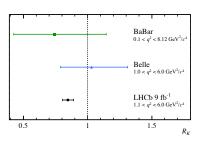


6 / 27

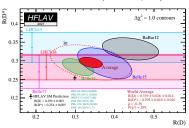
#### LFU Anomalies until few months ago

- Several R-like experiments suggested deviations from SM predictions
- More than  $3\sigma$  tension for both channels!

$$b\to sl^+l^-$$
 ex.  $R_K\propto \frac{\mathcal{B}(B^+\to K^+\mu^+\mu^-)}{\mathcal{B}(B^+\to K^+e^+e^-)}$ 



ex. 
$$R(D) = \frac{\mathcal{B}(B^+ \to D^0 \tau^+ \nu)}{\mathcal{B}(B^+ \to D^0 \mu^+ \nu)}$$



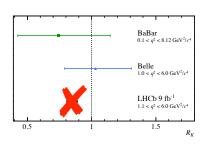
HFLAV 2021 Update

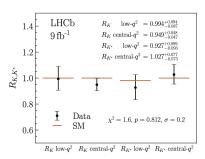
7 / 27

LHCb Nature Phys. 18 (2022) 277

#### **LFU Anomalies Now** - $b \rightarrow sl^+l^-$ **channel**

- A new LHCb measurement of  $R_K$  and  $R_K^*$  has been recently published, and it corrects the old published result.
- The new results are no more in tension with the SM prediction.





LHCb Nature Phys. 18 (2022) 277

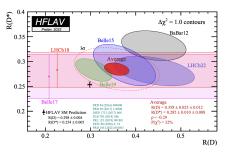
arXiv:2212.09153 arXiv:2212.09152

8 / 27

F.Riti  $R(J/\psi)$  Analysis 26 Jan 2023

#### LFU Anomalies Now - $b \rightarrow cl\nu$ channel

 A recent LHCb measurement of the combined channels of R(D) and R(D\*) confirmed the tension with the SM prediction in this channel!



Seminar

Therefore, even though things have changed for the  $b \to sl^+l^-$  channel, the  $b \to cl\nu$  one still looks very interesting and the LFU anomaly is still very much alive!

9 / 27

### $R(J/\psi)$ Analysis - a LFUV measurement in CMS

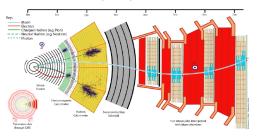
Another interesting R-measurement in the  $b \to c l \nu$  channel is the  $R(J/\psi)$  measurement  $\to$  **my PhD project**!

$$R(J/\psi) = \frac{\mathcal{B}(B_c^+ \to J/\psi \tau^+ \nu_\tau)}{\mathcal{B}(B_c^+ \to J/\psi \mu^+ \nu_\mu)}$$

- SM prediction:  $R(J/\psi) = 0.25 0.28$ , depending on the choice of form factors (Private Note G. Isidori);
- Only one previous result: LHCb experiment (Run I: 3  $fb^{-1}$ )  $R(J/\psi) = 0.71 \pm 0.17(stat) \pm 0.18(syst) = 0.71 \pm 0.25$   $\sim 2\sigma \text{ from SM prediction}$  (enhanced  $\tau$  couplings)

# $R(J/\psi)$ Analysis in CMS

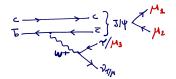
 The Compact Muon Solenoid (CMS) is a general-purpose detector at the Large Hadron Collider (LHC).



• My PhD project focuses in measuring  $R(J/\psi)$  from proton-proton collision data collected by the CMS experiment at LHC in 2018.

# $R(J/\psi)$ Analysis in CMS - Motivations

$$R(J/\psi) = \frac{\mathcal{B}(B_c^+ \to J/\psi \tau^+ (\to \mu^+ \bar{\nu_\mu} \nu_\tau) \nu_\tau)}{\mathcal{B}(B_c^+ \to J/\psi \mu^+ \nu_\mu)}$$



#### Why CMS?

- The  $B_c$  meson cannot be produced at the B-factories (Belle, BaBar, Bellell):  $e^+e^-$  at c.o.m energies around the Y(4S) peak. It can only be produced at hadron colliders ( $B_c$  was discovered at Tevatron (Phys. Rev. Lett. 81, 2432));
- $J/\psi \rightarrow \mu\mu$  can be very efficiently triggered;
- Higher luminosity than LHCb, , 59fb<sup>-1</sup> for 2018 data;
- For muonic  $\tau$  decay  $(\tau^+ \to \mu^+ \bar{\nu_\tau} \nu_\mu)$  no need to have particle ID detectors to tell hadrons apart, because there are 3 muons in the final state.

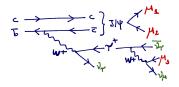
F.Riti  $R(J/\psi)$  Analysis

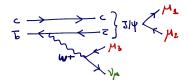
#### **Final State**

$$R(J/\psi) = \frac{\mathcal{B}(B_c^+ \to J/\psi \tau^+ (\to \mu^+ \bar{\nu_\mu} \nu_\tau) \nu_\tau)}{\mathcal{B}(B_c^+ \to J/\psi \mu^+ \nu_\mu)}$$

Num:  $B_c^+ \rightarrow J/\psi \tau^+ \nu_{\tau}$ 

Den:  $B_c^+ o J/\psi \mu^+ \nu_\mu$ 





13 / 27

- The only channel considered is the leptonic decay of  $\tau$ :  $\tau \to \mu^+ \bar{\nu_\mu} \nu_\tau$
- Therefore numerator and denominator have same visible final state of 3 muons and can be simultaneously reconstructed:
  - Two muons ( $\mu_1$  and  $\mu_2$ ) come from the  $J/\psi$ ;
  - The third muon  $\mu_3$  is displaced, and it either comes from the leptonic  $\tau$  decay (numerator), or directly from the  $B_c$  decay (denominator).

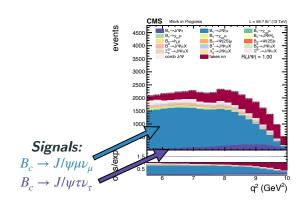
F.Riti  $R(J/\psi)$  Analysis 26 Jan 2023

#### **Decay kinematics: Neutrinos**

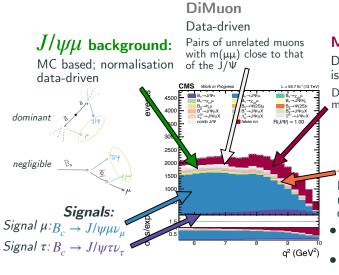
One of the challenges of this analysis is the presence of **neutrinos** in the final state:

- The  $B_c^{visible}$  candidate that can be reconstructed from the three muons in the final state is not the actual  $B_c^+$ , because neutrinos can not be directly detected.
  - To reconstruct the  $p_{B_c}$  we use the collinear approximation:  $B_c$  has the same direction of the visible final state, and  $p_{B_c} = \frac{m_B}{m_{pero}} p_{B_c}^{reco}$ .
- The analysis aims to separate  $3\nu$  (num.) vs  $1\nu$  (den.) decays leveraging these kinematical observables:
  - $q^2 = (p_B p_{J/\psi})^2$ ;
  - $m_{miss}^2 = (p_B p_{J/\psi} p_{\mu_3})^2$ ;
  - $p_T^{miss} = (p_T^B p_T^{J/\psi} p_T^{\mu_3});$
  - $E_{\mu}^*$ :  $\mu_3$  energy in the  $B_c$  candidate rest frame;
  - $E_{\mu}^{\#}$ : the  $\mu_3$  energy in the rest frame of the  $J/\psi$ .

### **Signal Contributions**



# Signal Contributions and Background Sources



#### Muon fakes

Data-driven antiisolated  $\mu$  sideband

Displaced  $J/\Psi$  + misidentified hadron

# $B_{c}$ background:

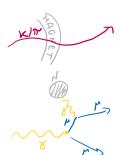
MC based; normalisation datadriven

- Feeddowns: excited  $c\bar{c}$  states to  $J/\psi$
- Other  $J/\psi$ +charmed hadron, mostly  $B_c \to D_s^{(*)} J/\psi$

#### Muon Fakes Background

- In the detector there are events where a  $J/\psi$  meson is coupled with objects coming from:
  - Decay in flight  $(K \to \mu \nu)$
  - Punch-through (hadrons that pass the magnet)
  - Photon conversion  $(\gamma \to \mu \mu)$
  - Actual fakes coming from accidental reconstruction
- Some of these object could be misidentified as muons → muon fakes.

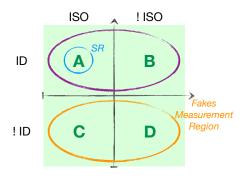




A key point of this analysis is to estimate this contribution.

#### Muon Fakes Background - Data Driven Method

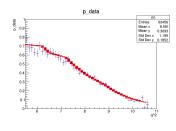
- Four regions are defined on  $\mu_3$  features: ID and isolation (ISO).
  - ID: how well the  $\mu$  is reconstructed;
  - ISO: how much energy is deposited around the  $\mu$ .
- The fakes bkg derivation is measured in the !ID region with NNs;
- After doing the measurement in the !ID region, the application is done in the ID region.



F.Riti

# Fakerate parametrisation using NNs: the Universal Approximation Theorem

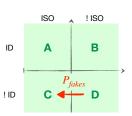
- A Neural Network (NN) can approximate any arbitrary complex f(x)
- This can be simply visualised in a 1D space
- But this is particularly useful if we need to parametrise for many variables!



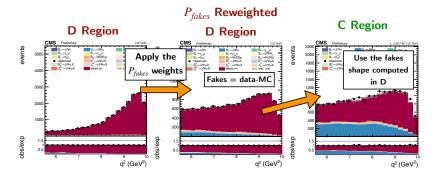
In our case: we need to find a transfer function  $P_{fakes}$  from !ISO to ISO regions in N dimensions  $\rightarrow$  we use a NN.

# Muon Fakes Background - Measurement of $P_{fakes}$

In the !ID region , event-by-event weights  $(P_{fakes})$  are computed with a NN, to account for the fake rate probabilities from !ISO to ISO regions.



 $Fakes(C) = P_{fakes} \cdot Data(D) - P_{fakes} \cdot MC(D)$ 

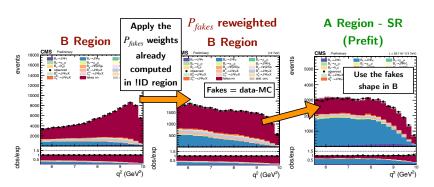


F.Riti  $R(J/\psi)$  Analysis

### Muon Fakes Background - Application to the Signal Region

P<sub>fakes</sub> weights, computed with the NNs already trained in the !ID region, can be applied to the events in the ID region, to find the fakes shape in A

$$Fakes(A) = P_{fakes} \cdot Data(B) - P_{fakes} \cdot MC(B),$$



F.Riti

#### Fit Model

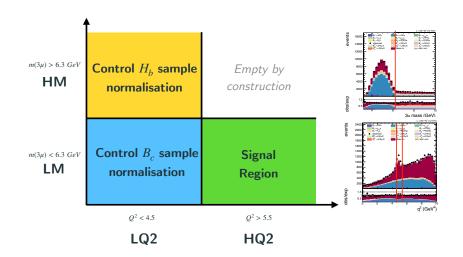
#### Some Challenges:

- The muon fakes background is estimated from data
- ullet The production cross sections  $\sigma$  are not well known, including  $B_c$   $\sigma$

#### The fit model:

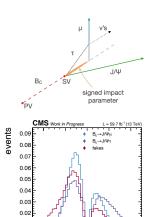
- · Binned maximum likelihood fit
- The result of the fit is directly the  $R(J/\psi)$  value
- The normalisations of the produced samples are constrained thanks to control regions
- The fakes estimate is in-situ as part of the simultaneous fit
- All the uncertainties of the fit are included as either rate or shape uncertainties

# Fit Model - Simultaneous to Multiple Categories



# Fit Model - Improve Sensitivity

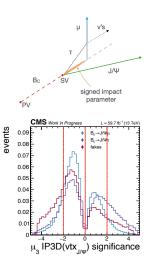
- The signed impact parameter 3D variable (IP3D) is defined as the distance 3D of  $\mu_3$  from the  $J/\psi$  vertex
- The  $IP3D_{sig} = \frac{IP3D}{IP3D_{err}}$  variable shows a different behaviour for fakes bkg and the two signals, therefore it is useful to improve the sensitivity of the analysis: we can find a phase space where the S/B ratio is better.



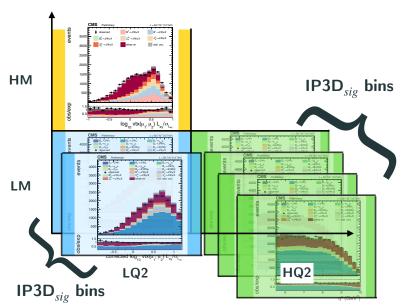
0.0

# Fit Model - Improve Sensitivity

- The signed impact parameter 3D variable (IP3D) is defined as the distance in z of  $\mu_3$  from the  $J/\psi$  vertex
- The  $IP3D_{sig} = \frac{IP3D}{IP3D_{err}}$  variable shows a different behaviour for fakes bkg and the two signals, therefore it is useful to improve the sensitivity of the analysis: we can find a phase space where the S/B ratio is better.



# Fit Model Summary - 14 categories: 7 (ISO) & 7 (!ISO))



24 / 27

#### **Systematic Uncertainties**

#### The most important systematics are:

- Form Factors to update the old models used in the simulations with new theories → they impact a lot q<sup>2</sup> shape;
- Several fakes background estimation uncertainties;
- BR and normalisation uncertainties on the BRs of background processes with genuine muons;
- Experimental Uncertainties:  $\mu$  features' corrections; pileup corrections;  $B_c$  lifetime.

#### Sensitivity of the analysis

- We performed an Asimov fit, to find the sensitivity of the analysis;
- We use as expected signal the LHCb result;
- The result is  $r = 0.71 \pm ??$

The sensitivity is embargoed because the analysis is still under review!

#### **Conclusions**

- Searches for LFUV are a very hot topic in particle physics!
- CMS has the sensitivity to play an important role on this topic
- $R(J/\psi)$  is a very interesting measurement to look for hints of new physics, only possible at LHC
- It is a challenging measurement to perform!

This analysis is amongst the firsts LFUV measurements performed at CMS, and it paves the way for further measurements.

Backup

#### Preselection

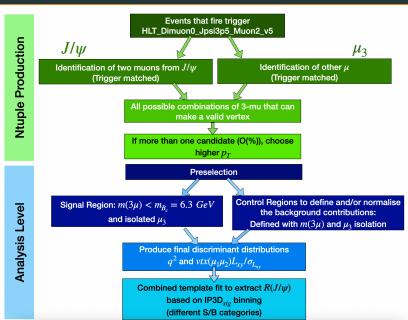
#### Final State particles:

- $\mu_1$ : mediumID,  $p_T^{\mu_1} >$  6,  $|\eta^{\mu_1}| <$  2.5,  $|d_{xy}^{\mu_1}| <$  0.05
- $\mu_2$ : mediumID,  $p_T^{\mu_2} >$  4,  $|\eta^{\mu_2}| <$  2.5,  $|d_{xy}^{\mu_2}| <$  0.05
- $\mu$ :  $p_T^{\mu} > 4$ ,  $|\eta^{\mu}| < 2.5$ ,  $|d_{xy}^{\mu}| < 0.05$
- $|d_z^{\mu_1} d_z^{\mu_2}| < 0.2$ ,  $|d_z^{\mu_1} d_z^{\mu}| < 0.2$ ,  $|d_z^{\mu_2} d_z^{\mu}| < 0.2$
- $\Delta R_{12} > 0.01, \Delta R_{13} > 0.01, \Delta R_{23} > 0.01$

#### $J/\psi$ and $B_c$ vertices properties:

- $prob_{J/\psi} > 0.01$
- $prob_{B_c} > 10^{-4}$
- $|m^{J/\psi} 3.0969| < 0.1$

#### **Analysis Strategy**



F.Riti

 $R(J/\psi)$  Analysis

# Analytical derivation for fakes shape in C (1/2)

$$Data(C) = MC(C) + Fakes(C)$$
 (1)

multiplying and dividing by the respective shapes in D, we obtain

$$\frac{Data(C)}{Data(D)} \cdot Data(D) = \frac{MC(C)}{MC(D)} \cdot MC(D) + \frac{Fakes(C)}{Fakes(D)} \cdot Fakes(D). \quad (2)$$

The ratios can be defined as

$$\frac{Data(C)}{Data(D)} = P_{data}; \ \frac{MC(C)}{MC(D)} = P_{MC}; \ \frac{Fakes(C)}{Fakes(D)} = P_{fakes}; \ \frac{MC(D)}{Data(D)} = \alpha$$
(3)

N.B. these are not actual probabilities therefore Eq.2 can be then rewritten as

$$P_{data} \cdot Data(D) = P_{MC} \cdot MC(D) + P_{fakes} \cdot Fakes(D)$$
 (4)

 $R(J/\psi)$  Analysis 26 Jan 2023

# Analytical derivation for fakes shape in C (2/2)

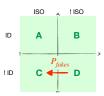
$$P_{data} \cdot Data(D) = P_{MC} \cdot MC(D) + P_{fakes} \cdot Fakes(D)$$

Considering Fakes(D) = Data(D) - MC(D) and dividing everything by Data(D), the following equation is obtained:

$$P_{data} = P_{MC} \cdot \alpha + P_{fakes} \cdot (1 - \alpha) \tag{5}$$

Solving for  $P_{fakes}$ :

$$P_{fakes} = \frac{P_{data} - P_{MC} \cdot \alpha}{(1 - \alpha)}. \tag{6}$$



 $P_{fakes} = \frac{Fakes(C)}{Fakes(D)}$  , then the shape of the fakes background in the C region

$$\textit{Fakes}(\textit{C}) = \frac{\textit{P}_{\textit{data}}(\textit{x}_{\textit{i}}) - \textit{P}_{\textit{MC}}(\textit{x}_{\textit{i}}) \cdot \alpha(\textit{x}_{\textit{i}})}{1 - \alpha(\textit{x}_{\textit{i}})} \cdot \textit{Data}(\textit{D}) - \frac{\textit{P}_{\textit{data}}(\textit{x}_{\textit{i}}) - \textit{P}_{\textit{MC}}(\textit{x}_{\textit{i}}) \cdot \alpha(\textit{x}_{\textit{i}})}{1 - \alpha(\textit{x}_{\textit{i}})} \cdot \textit{MC}(\textit{D})$$

From now on the event-by-event dependency will be implicit

F.Riti R(J/

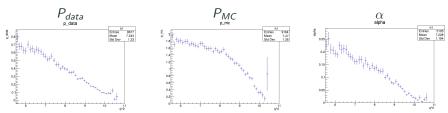
31 / 27

# Find $P_{data}$ , $P_{MC}$ and $\alpha$ as $q^2$ histos ratio

**How can we compute this?** The final formula depends on the three sets of weights:

$$P_{\textit{data}} = \frac{\textit{Data}(\textit{C})}{\textit{Data}(\textit{D})}; \quad P_{\textit{MC}} = \frac{\textit{MC}(\textit{C})}{\textit{MC}(\textit{D})}; \quad \alpha = \frac{\textit{MC}(\textit{D})}{\textit{Data}(\textit{D})}$$

These weights can be computed with bin per bin ratios of histos, for example in  $q^2$  variable.

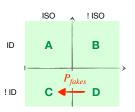


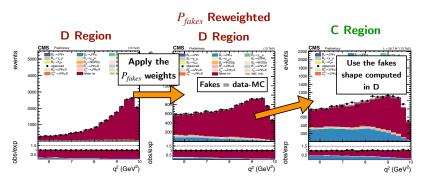
Once we have these 3 sets of weights, the final weight  $P_{\mathit{fakes}} = \frac{P_{\mathit{data}} - P_{\mathit{MC}} \cdot \alpha}{(1 - \alpha)} \text{ can be computed}.$ 

# Application of the $P_{fakes}$ weights

Once the weights  $P_{fakes} = \frac{P_{data} - P_{MC} \cdot \alpha}{(1-\alpha)}$  are computed with the 3 histos ratios, they are applied to the events in the D region, and the fakes shape in C is found as

$$Fakes(C) = P_{fakes} \cdot Data(D) - P_{fakes} \cdot MC(D)$$





### Fakerate parametrisation through NNs

Instead of computing the ratios of histos, these weights can be found using NNs:

- 1. To have a continuous output;
- 2. To be able to parametrize for more than one variable.

Use classification neural networks and interpreting the outputs as weights.

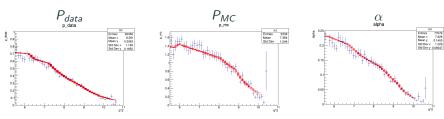
If we train 3 classification NNs, we can obtain the 3 sets of weights that we need to compute  $P_{fakes}$  (ref.) :

- P<sub>data</sub> → Data(C) vs Data(D);
- $P_{MC} \rightarrow MC(D)$  vs MC(C);
- $\alpha \rightarrow MC(D)$  vs Data(D).

# Validation of NN method using only $q^2$

When the NNs are trained with just  $q^2$ , the resulting weights should be equivalent to the ratios of histograms shown before.

In the plots, the points in red are the outputs of the NNs, and in blue the ratio of  $q^2$  histos (already shown in previous slide).



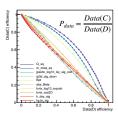
The two methods are equivalent if we use just one feature, therefore we can proceed using the NNs with more features.

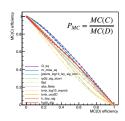
# Defining the input for the 3 NNs 1/2

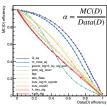
- P<sub>fakes</sub> are event-per-event weights, function of P<sub>data</sub>, P<sub>MC</sub> and α.
   Therefore 3 classification NNs are trained using the most discriminating features for the three;
- These features were chosen with the ROC curves pre training: for each feature, a ROC is built, and the distance from the diagonal indicates how much the two terms of the classification are different, and therefore how much the feature is useful in the training.
- Only features that are not correlated are kept in the training.

# Defining the input for the 3 NNs 2/2

- $\bullet$   $Q^2$
- |η<sub>B</sub>|
- $p_T^B$
- $m_{miss}^2$
- $log_{10}vtx(\mu_1,\mu_2)L_{xy}/\sigma_{L_{xy}}$
- $vtx(\mu_1, \mu_2, \mu_3)$  prob
- $(\mu_1, \mu_2, \mu_3) 2 D \cos \alpha$
- $\mu_3 IP3D(vtx_{J/\Psi})$  sig
- $\mu_3 |d_{xy}|/\sigma_{d_{xy}}$
- $\mu_3 |d_z|/\sigma_{d_z}$







#### Training the 3 NNs - Correlation Matrix

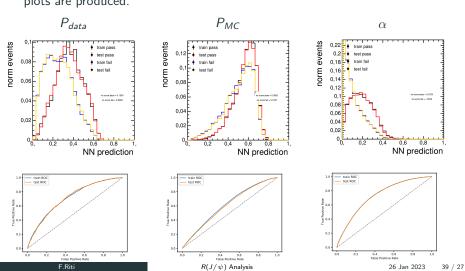


 $R(J/\psi)$  Analysis

1.00 0.75 - 0.50 - 0.25 - 0.00 - -0.25 - -0.50 - -0.75 -1.00

#### Training the 3 NNs - diagnostics

It is very important to check for overfitting of the networks, because the method relies on their ability to generalization  $\rightarrow$  therefore diagnostics plots are produced.



#### Validation of the method

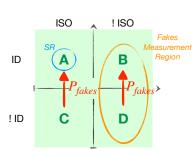
This method has been validated with some studies (see backup for the details):

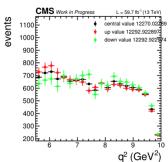
- 1. Closure test in the C region;
- 2. Check the extrapolation to the A region (SR) using the fakes MC  $H_b \rightarrow J/\psi X$ , also through the construction of a toy model.

#### Uncertainty on the method

How to put an uncertainty on the method?

- "Rotate" the method and compare the final shapes.
- The final fakes distributions in A differ in normalisation (already taken into account in the fit with a norm uncertainty) and shape
  - The nominal fakes shape is taken to be the weighted average of the two methods (bcs the rotated method suffer of higher stat unc)
  - Up and Down unc. in the fit are respectively the original and rotated methods





#### Fit details - Ratio



- Bin-by-bin: MC statistics;
- Form Factors ;
- BR and normalisation: uncertainties on the BRs in MC forced decays;
- fakes background estimation uncertainty;
- Experimental Uncertainties: scale factors; pileup;  $B_c$  lifetime.

The idea of "ratio" is achieved yia correlating  $B_c \to J/\psi\mu\nu$  and  $B_c \to J/\psi\mu\nu$  and  $B_c \to J/\psi\mu\nu$  and  $B_c \to J/\psi\mu\nu$