Measuring the decay $B^+ \to \rho^0 \mu^+ \nu_\mu$ at LHCb

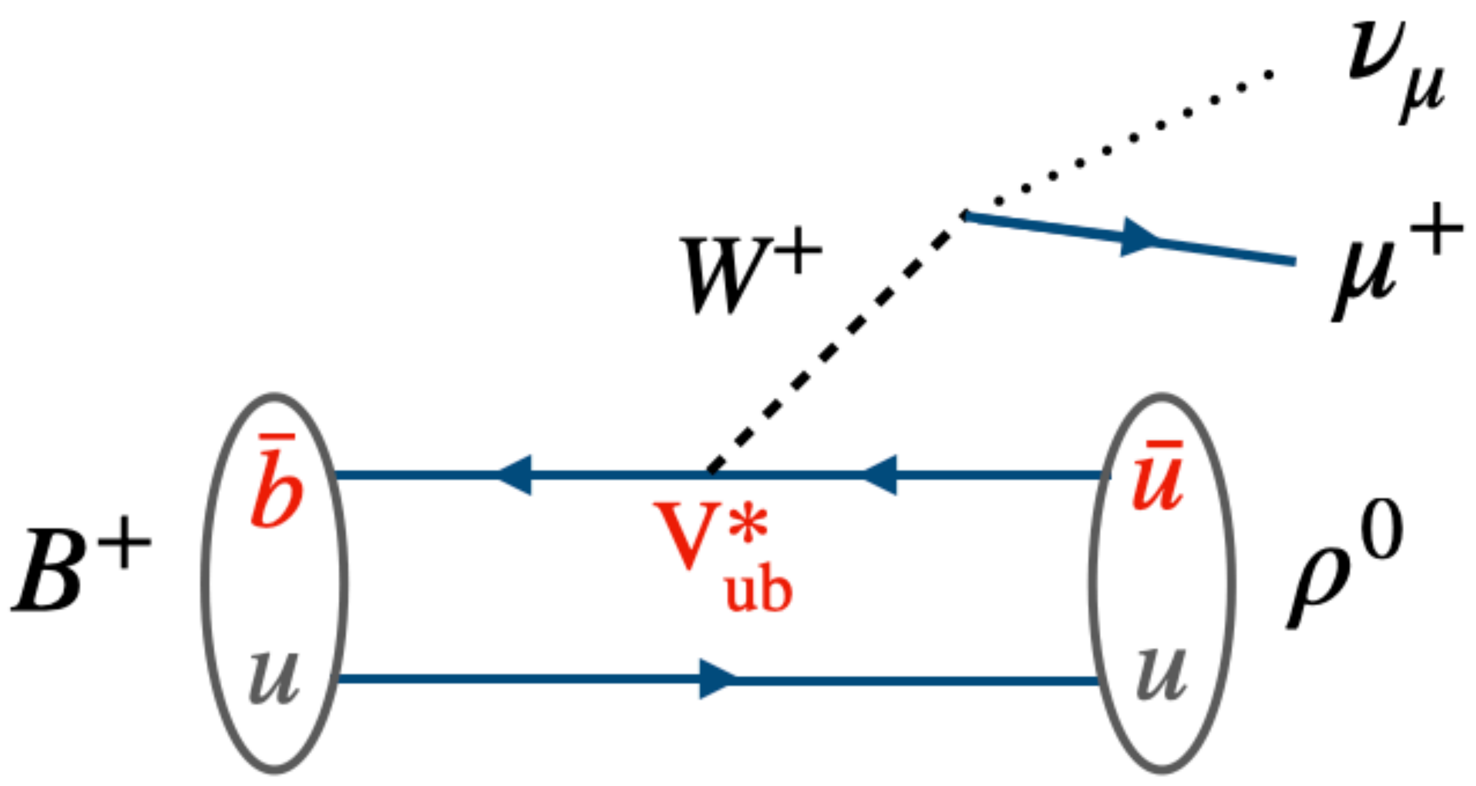
Veronica S. Kirsebom & Michel De Cian

EPFL

SPS, Innsbruck
September, 2021

What is $B^+ \to \rho^0 \mu^+ \nu_\mu$?

- **Semileptonic decay**:



CKM matrix

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.221 & 0.987 & 0.041 \\ 0.008 & 0.039 & 1.013 \end{pmatrix}$$

- **Belle collaboration** in 2013 [1]:

$\mathscr{B}(B^- \to \rho^0 l^- \bar{\nu}_l) = (1.83 \pm 0.20 \pm 0.10) \cdot 10^{-4}$ (with 621.7 ± 35.0 signal candidates)

- **Measuring $B^+ \to \rho^0 \mu^+ \nu_\mu$ at LHCb:**

Advantage: **~100 times more signal candidates**.

Challenges: hadronic environment, cannot precisely determine the momentum and number of produced B mesons.

[1] The Belle Collaboration, Phys. Rev. D 88, 032005 (2013).

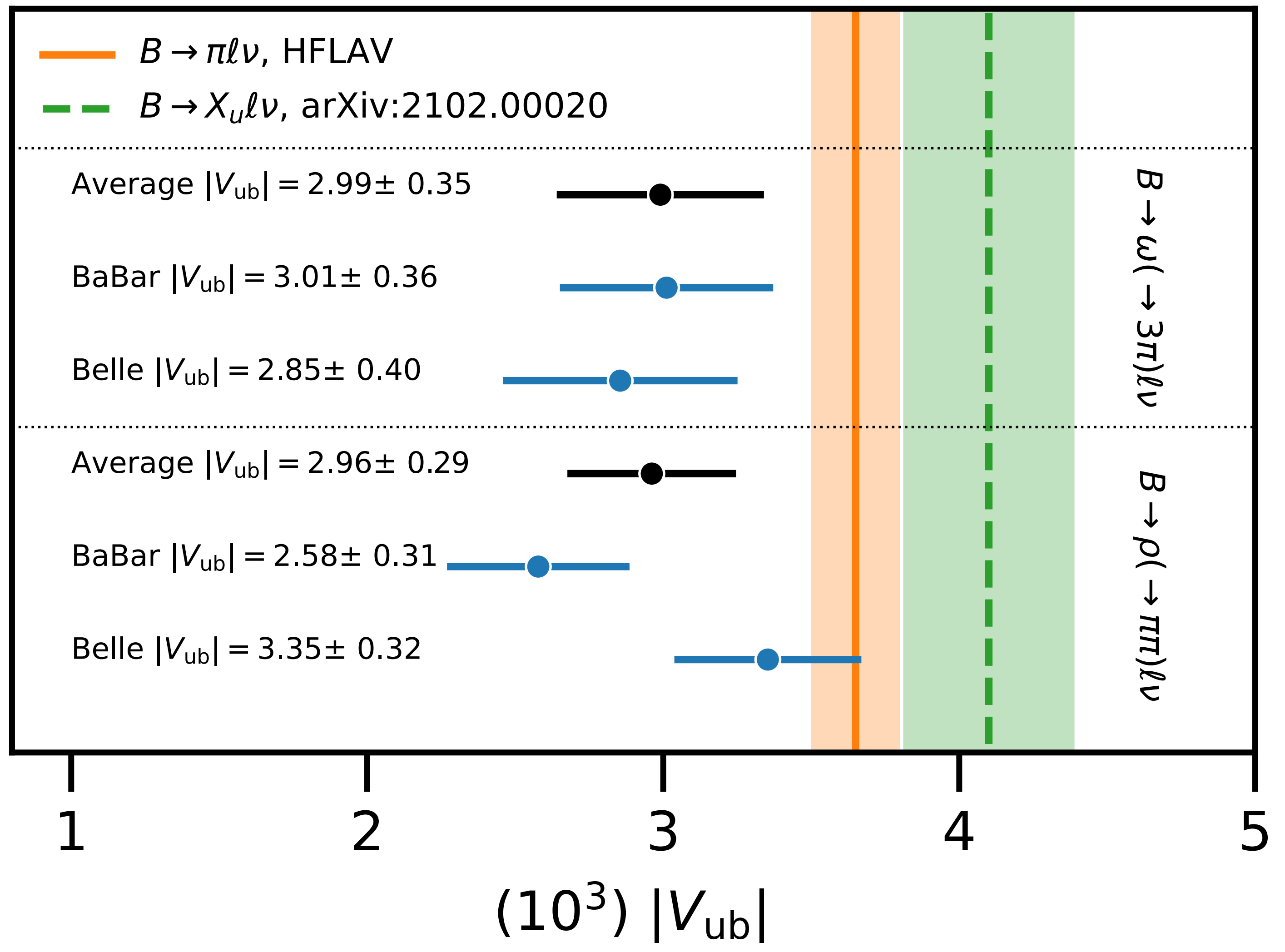
Why study $B^+ \to \rho^0 \mu^+ \nu_\mu$?

- Long-standing **tension between measurements of** $|V_{ub}|$ in **inclusive** and **exclusive** semileptonic decays.

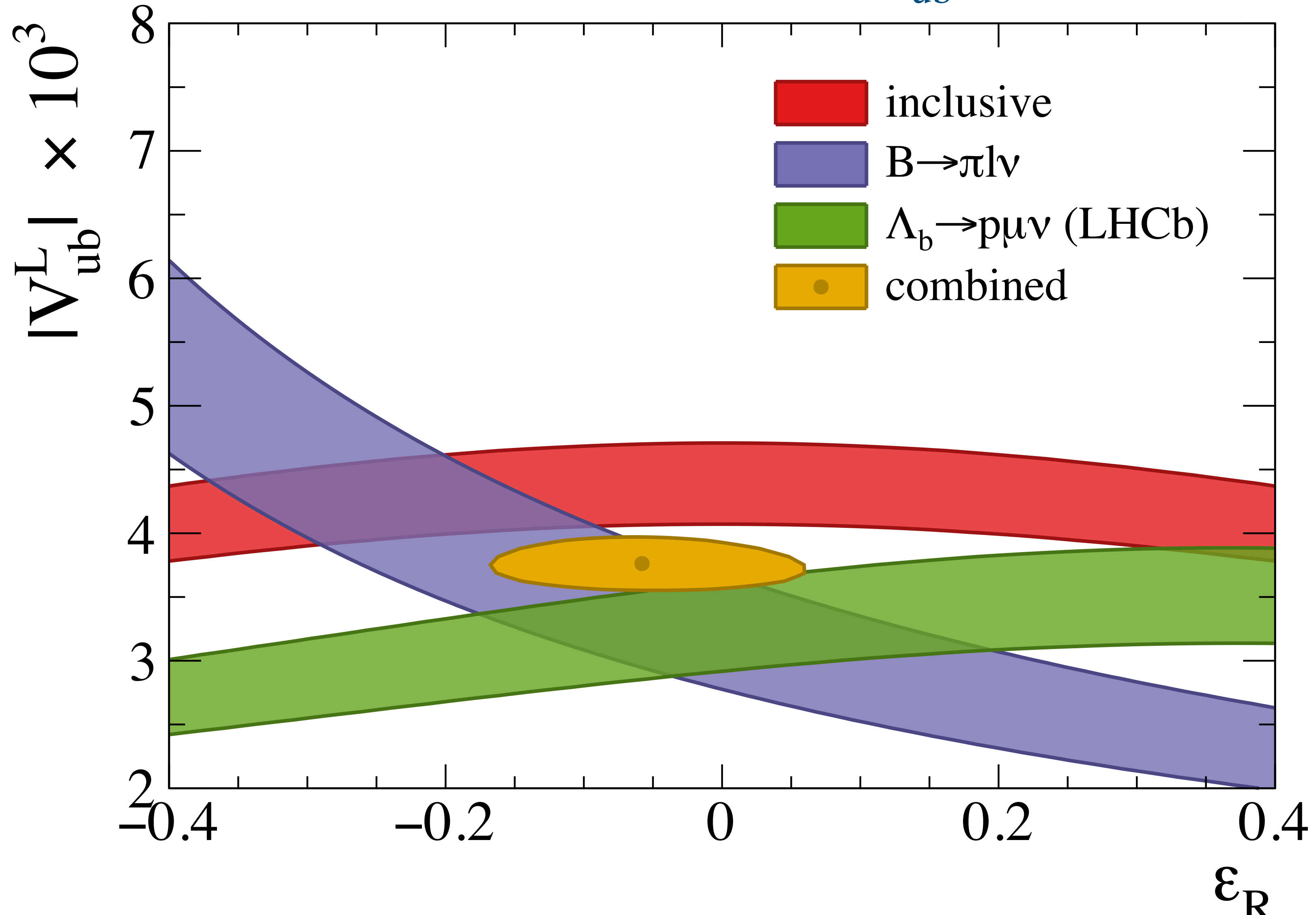
- Tension could be due to **new physics (NP)**, such as a right-handed weak current (parametrized with $\epsilon_R \neq 0$).

- **Measuring $B^+ \to \rho^0 \mu^+ \nu_\mu$ at LHCb can give us a more precise determination of $|V_{ub}|$ helping us to understand the $|V_{ub}|$ puzzle and a possible ϵ_R explanation.**

$|V_{ub}|$ determined from different channels [1]



Experimental constraints on $|V_{ub}^L|$ and ϵ_R [2]



[1] Florian U. Bernlochner et al. arXiv:2104.05739 (2021).
[2] The LHCb Collaboration, Nature Physics 11, 743–747 (2015).

Goals of this analysis

- **First goal:** Measure the differential decay rate as a function of q^2 and extract $|V_{\mathrm{ub}}|$.

$$\frac{\Gamma(B^+ \to \rho^0[\to \pi^+\pi^-]\mu^+\nu_\mu)}{dq^2}$$

Use $B^+ \to D^0[\to \pi^+\pi^-]\mu^+\nu_\mu$ as the normalisation channel.

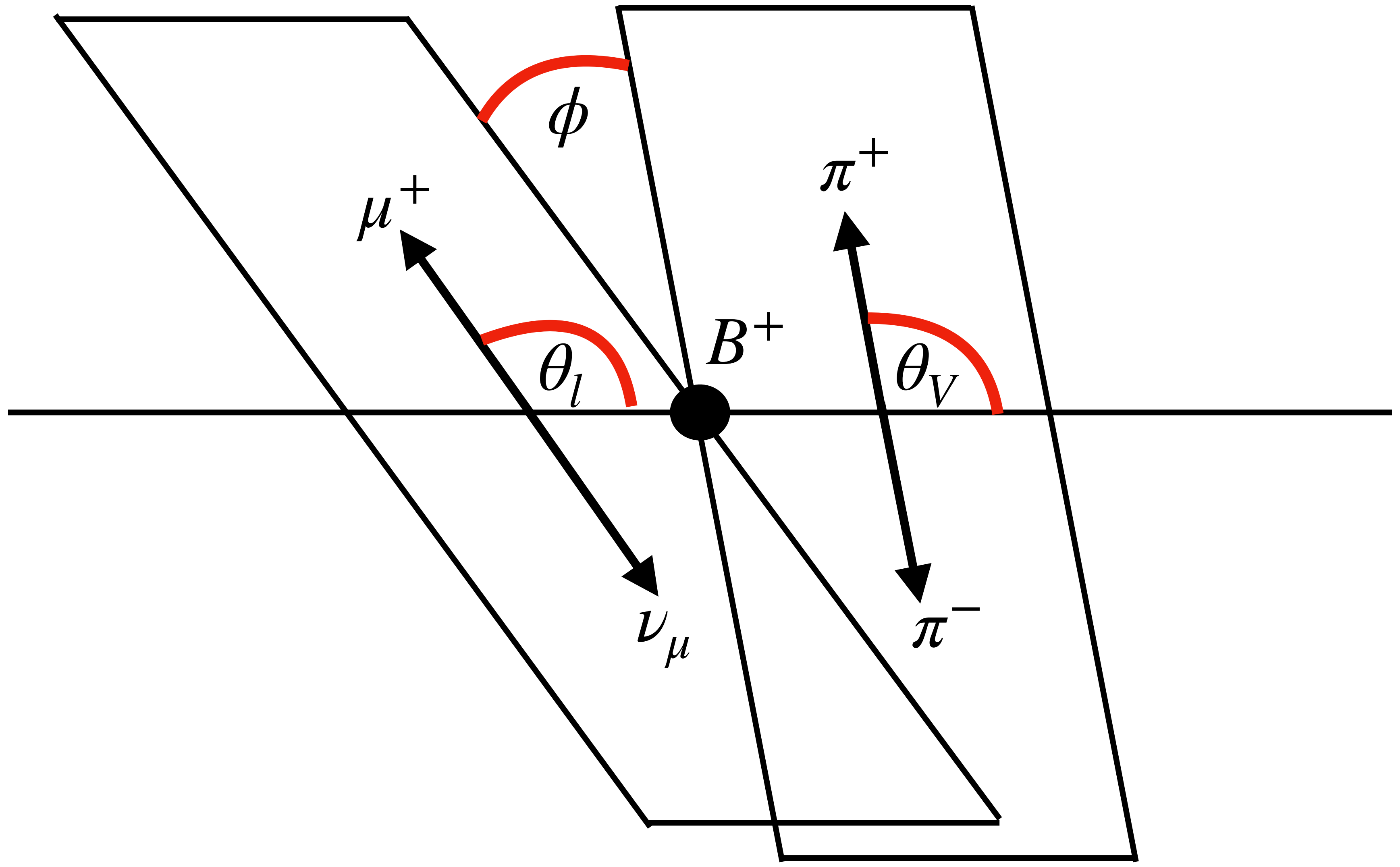
Focus of this talk

- **Final goal:** Measure the full differential decay rate as a function of q^2, θ_l, θ_V and ϕ, and extract ϵ_R.

Full differential decay rate [1]:

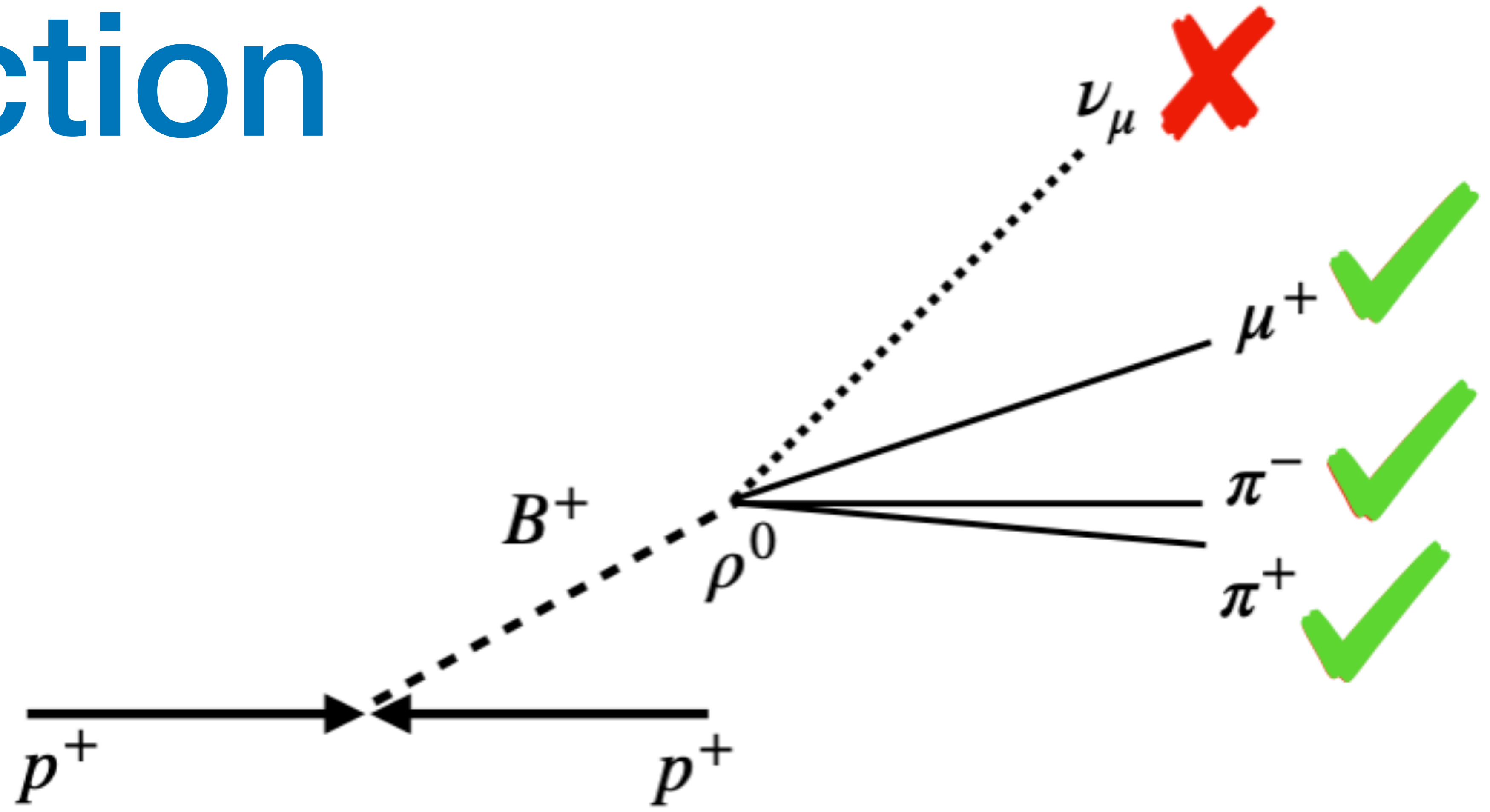
$$\frac{d\Gamma}{dq^2 d\cos\theta_V d\cos\theta_l d\phi} = \frac{G_F^2|V_{ub}|^2 m_B^3}{2\pi^4} \times \Big(J_{1s}\sin^2\theta_V + J_{1c}\cos^2\theta_V \\ + (J_{2s}\sin^2\theta_V + J_{2c}\cos^2\theta_V)\cos 2\theta_l + J_3\sin^2\theta_V\sin^2\theta_l\cos 2\phi \\ + J_4\sin 2\theta_V\sin 2\theta_l\cos\phi + J_5\sin 2\theta_V\sin\theta_l\cos\phi \\ + (J_{6s}\sin^2\theta_V + J_{6c}\cos^2\theta_V)\cos\theta_l + J_7\sin 2\theta_V\sin\theta_l\sin\phi \\ + J_8\sin 2\theta_V\sin 2\theta_l\sin\phi + J_9\sin^2\theta_V\sin^2\theta_l\sin 2\phi \Big).$$

Decay angles:



[1] Florian U. Bernlochner et al. Phys. Rev. D 90, 094003 (2014).

Signal reconstruction



- To reconstruct the signal, with an **unmeasured neutrino**, we use:

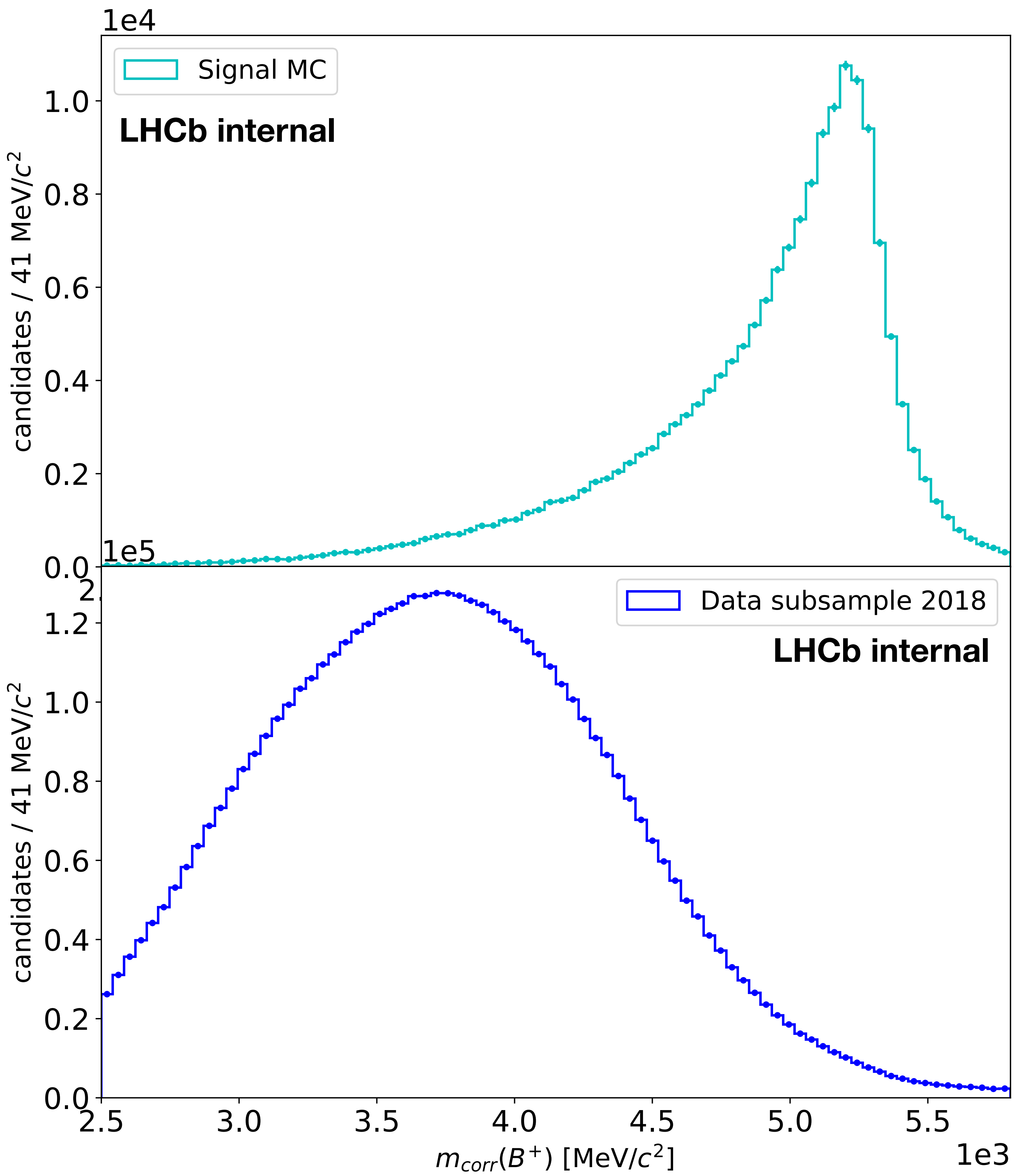
  Corrected B mass: $m_{corr}(B^+) = \sqrt{m_{vis}^2 + p_\perp^2} + p_\perp$

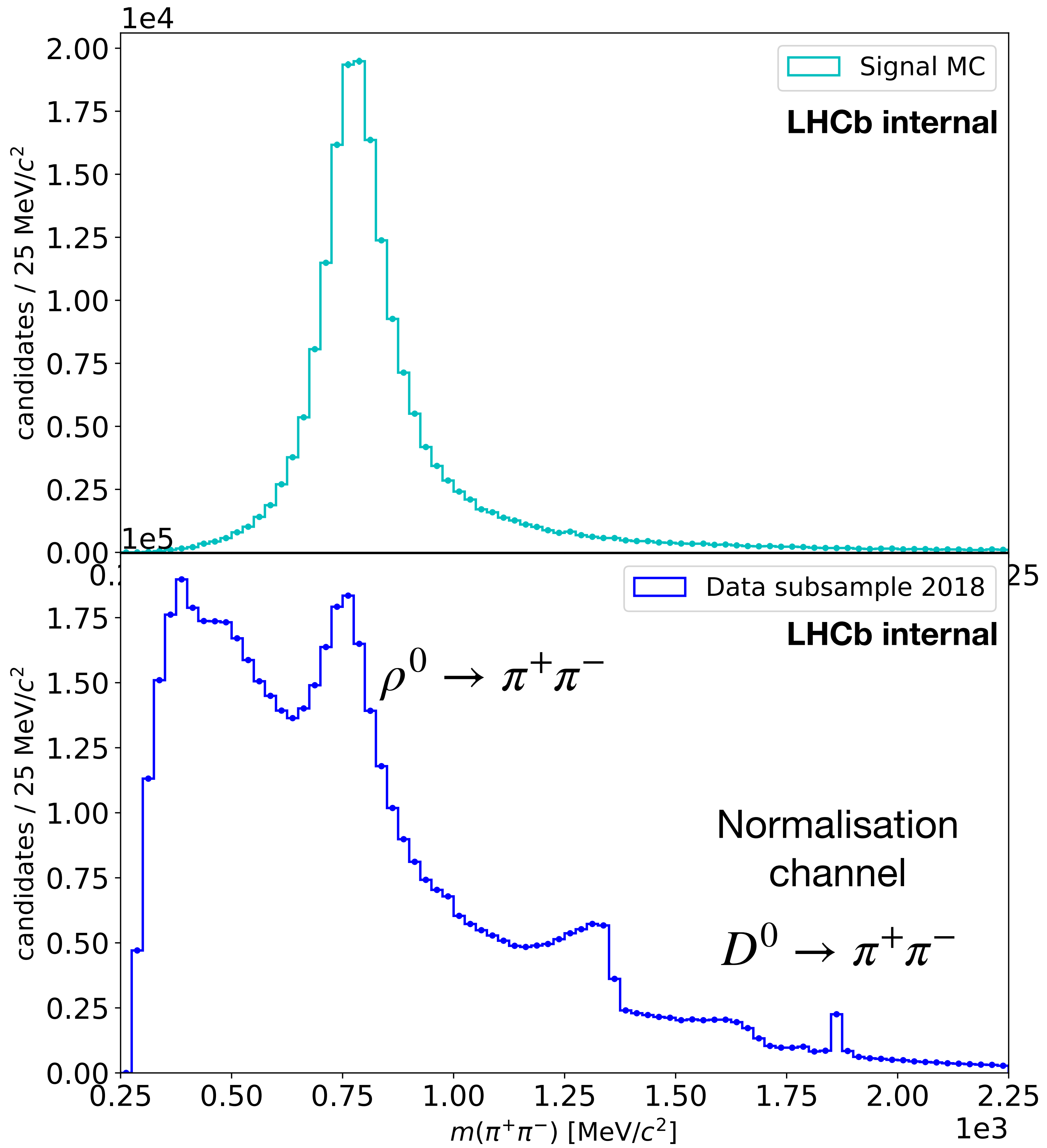
  The invariant mass of the two pions: $m(\pi^+\pi^-)$

- Plots of $m_{corr}(B^+)$ and $m(\pi^+\pi^-)$ **before** any selection cuts:

Signal-MC →

Data →

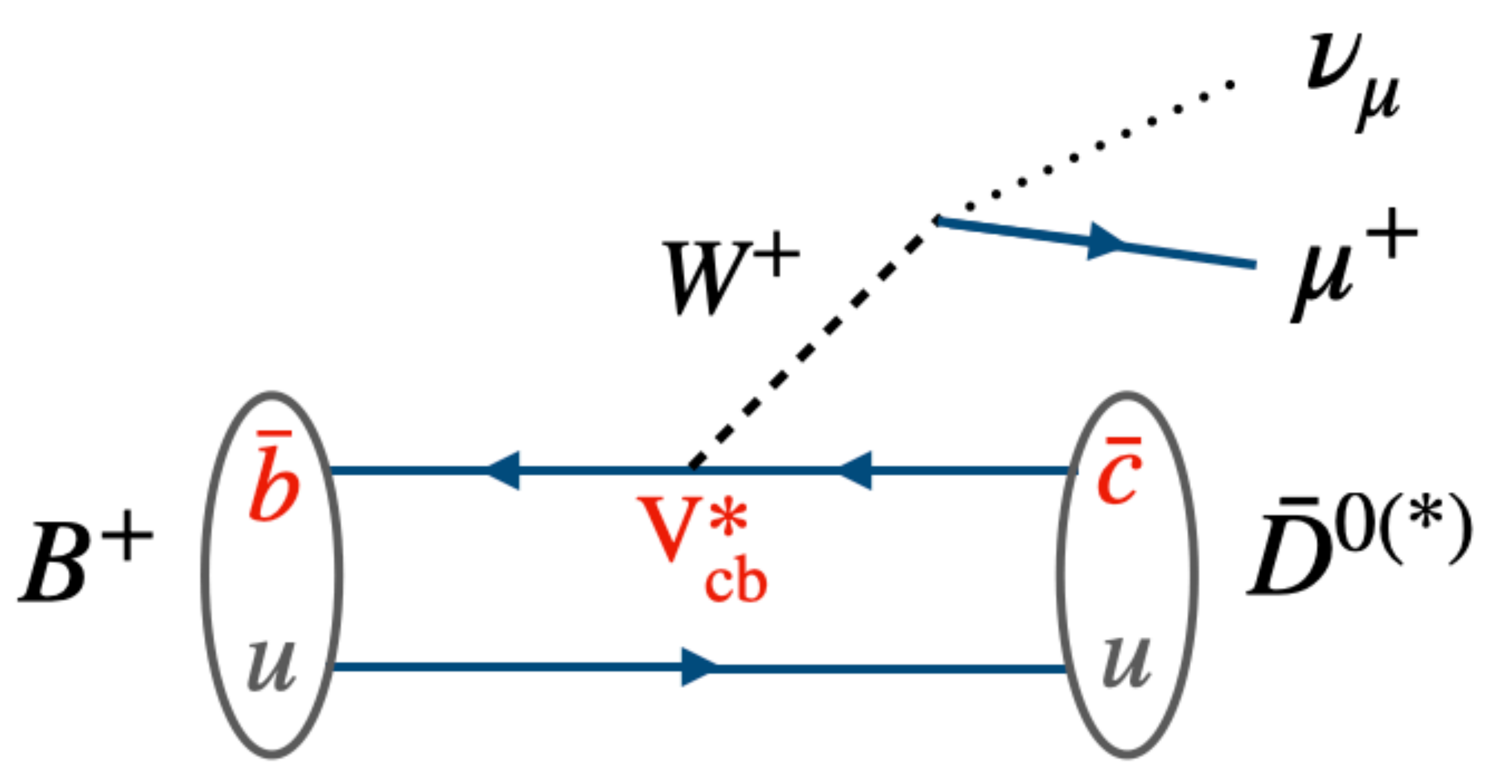




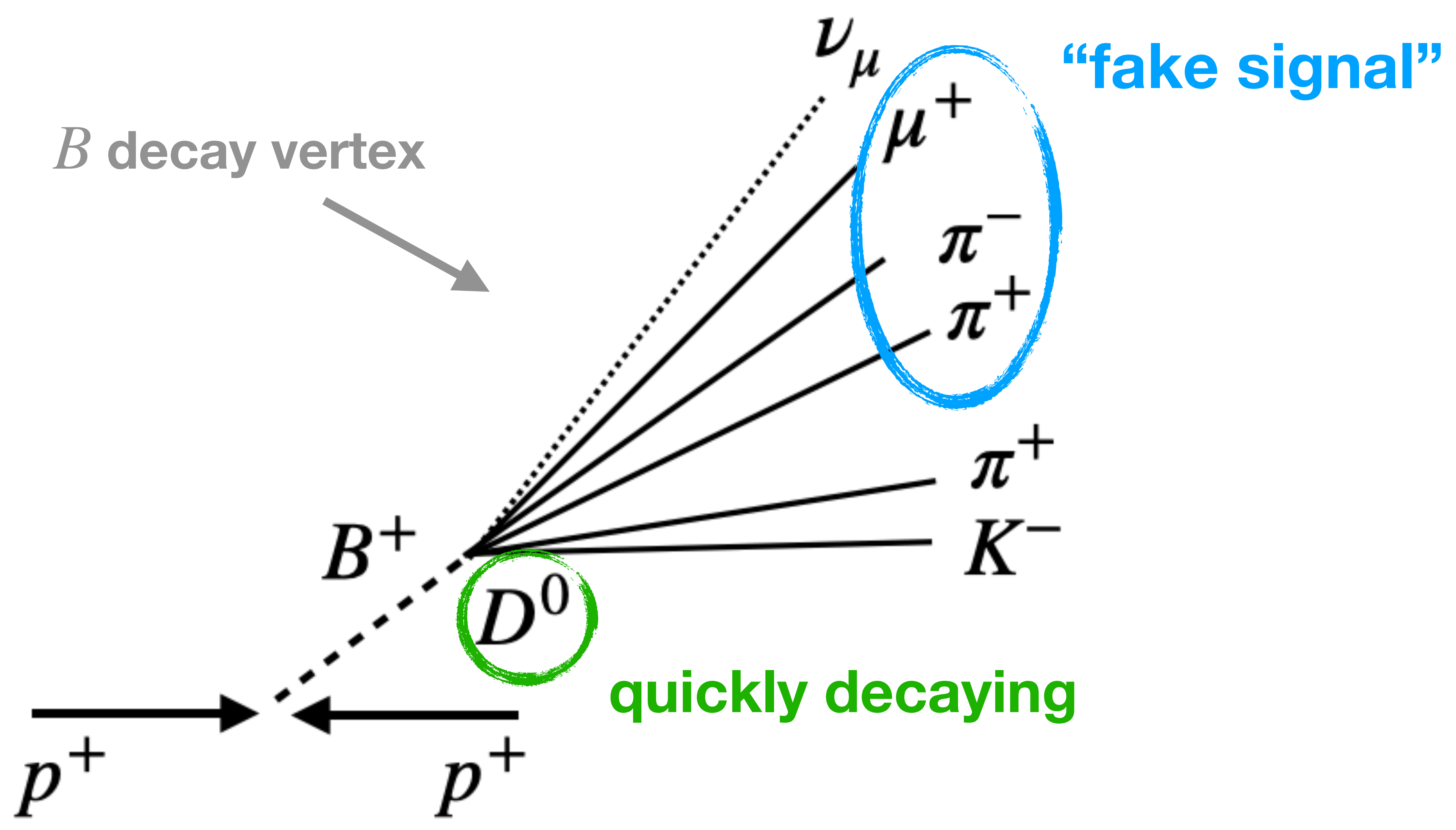
large background!

What are the backgrounds?

- **Dominant background:** **semileptonic decays** governed by the V_{cb} matrix element:



Example: $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^+\pi^-)\mu^+\nu_\mu$

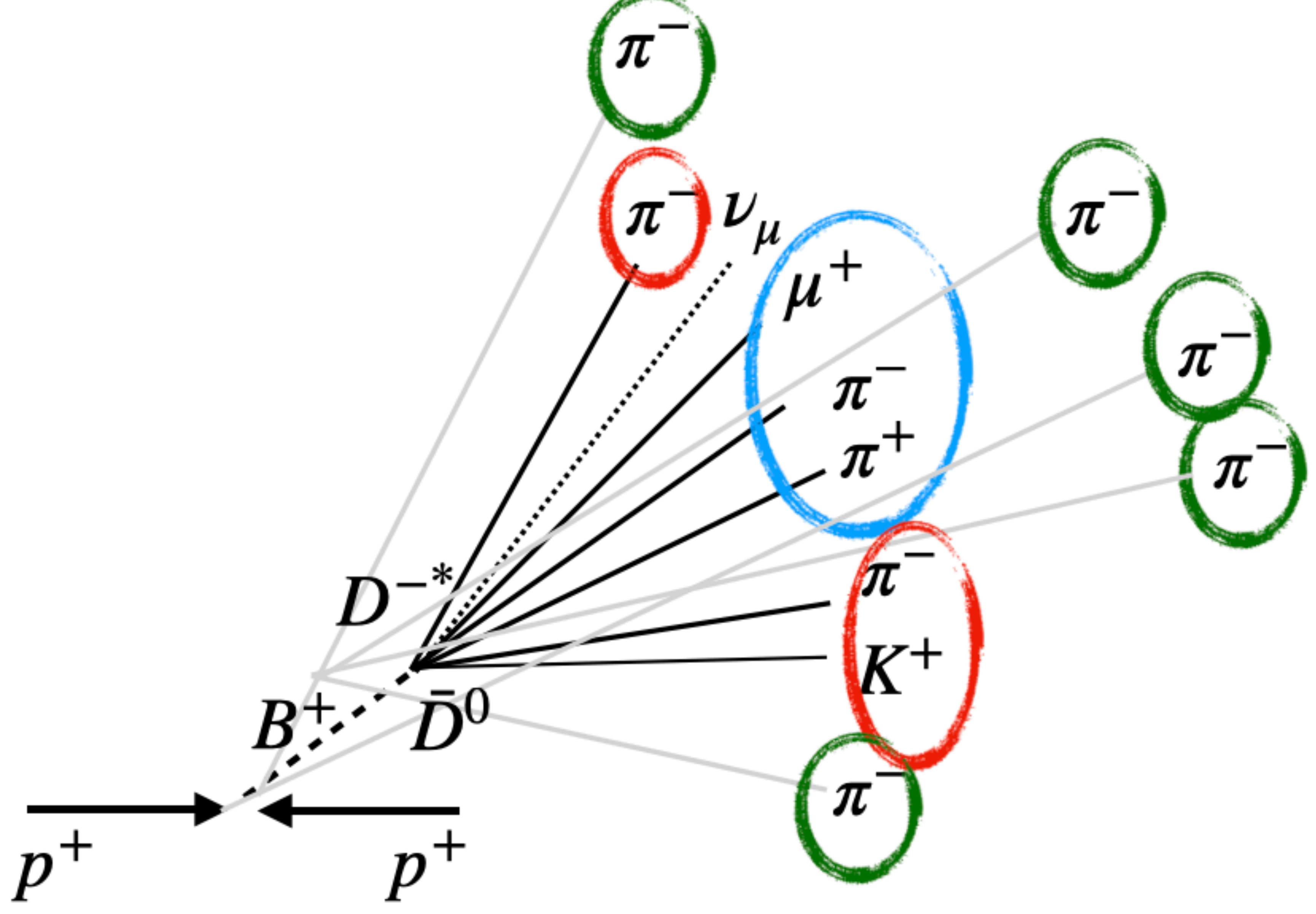


- This background is much **more abundant** than our signal:

$$\frac{\Gamma(b \to c)}{\Gamma(b \to u)} \propto \frac{|V_{cb}|^2}{|V_{ub}|^2} \approx 100$$

- **Other backgrounds:** $b \to u$ **semileptonic** decays and **combinatorial** background (combination of random tracks).

MVA selection



- Since our **most prominent backgrounds** often come with one or more **additional tracks**, we use an MVA trained on the **isolatedness** of the sig background.

- **How does it work?**

  1. We train an MVA to distinguish **tracks coming from the B decay chain** and **tracks coming from other processes**.

  2. We take our **signal candidate** and add, one by one, the other tracks in the event, and we use the **MVA to rank each track** according to how likely it is to come from the B decay chain**.**

  3. We add the **highest ranked tracks** to the **signal candidate**, and compute so-called “**isolation variables**”

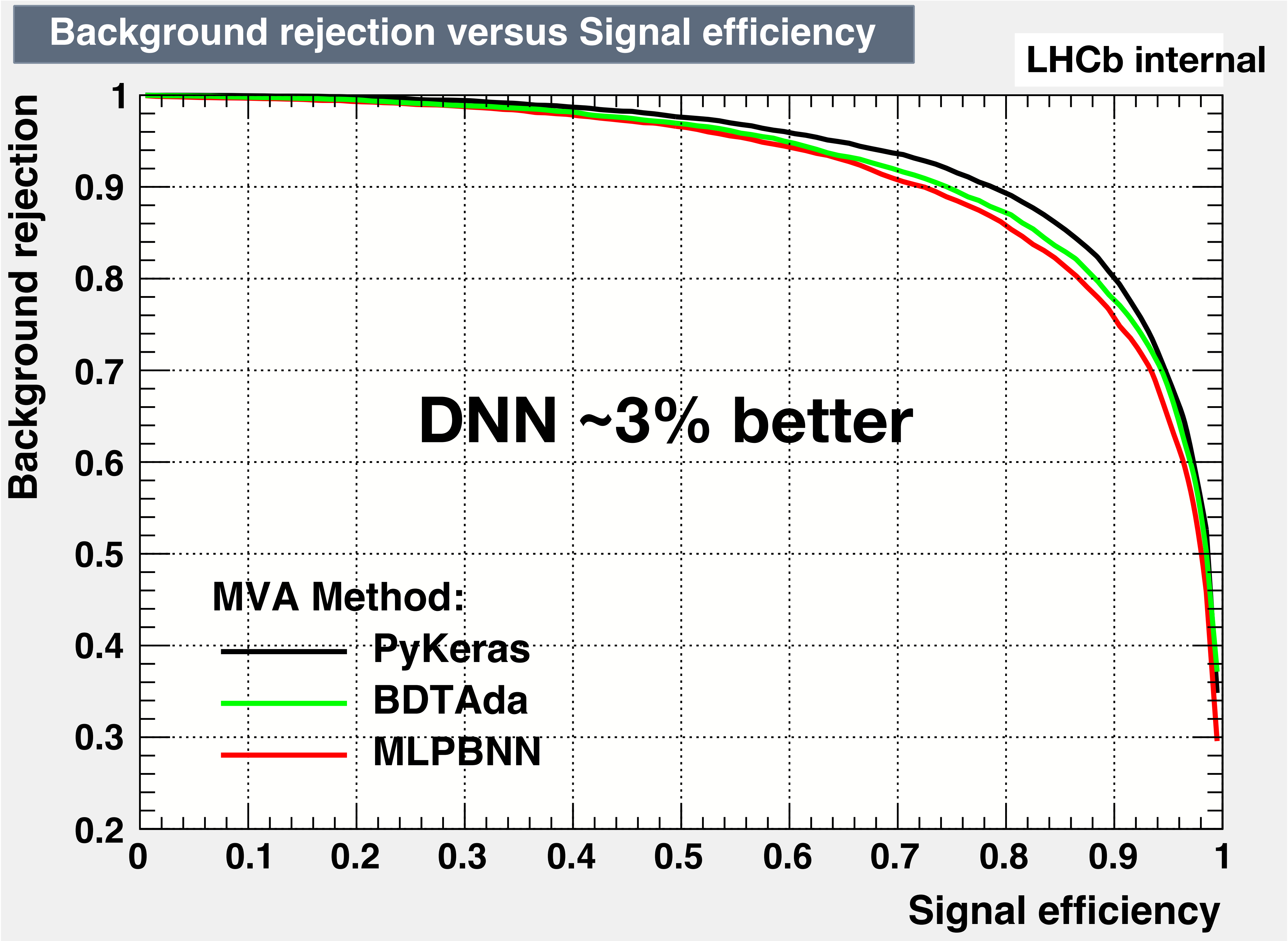
  4. We train a final MVA, deep neural network (DNN), with **isolation variables** (plus kinematic and geometric variables).

Examples:

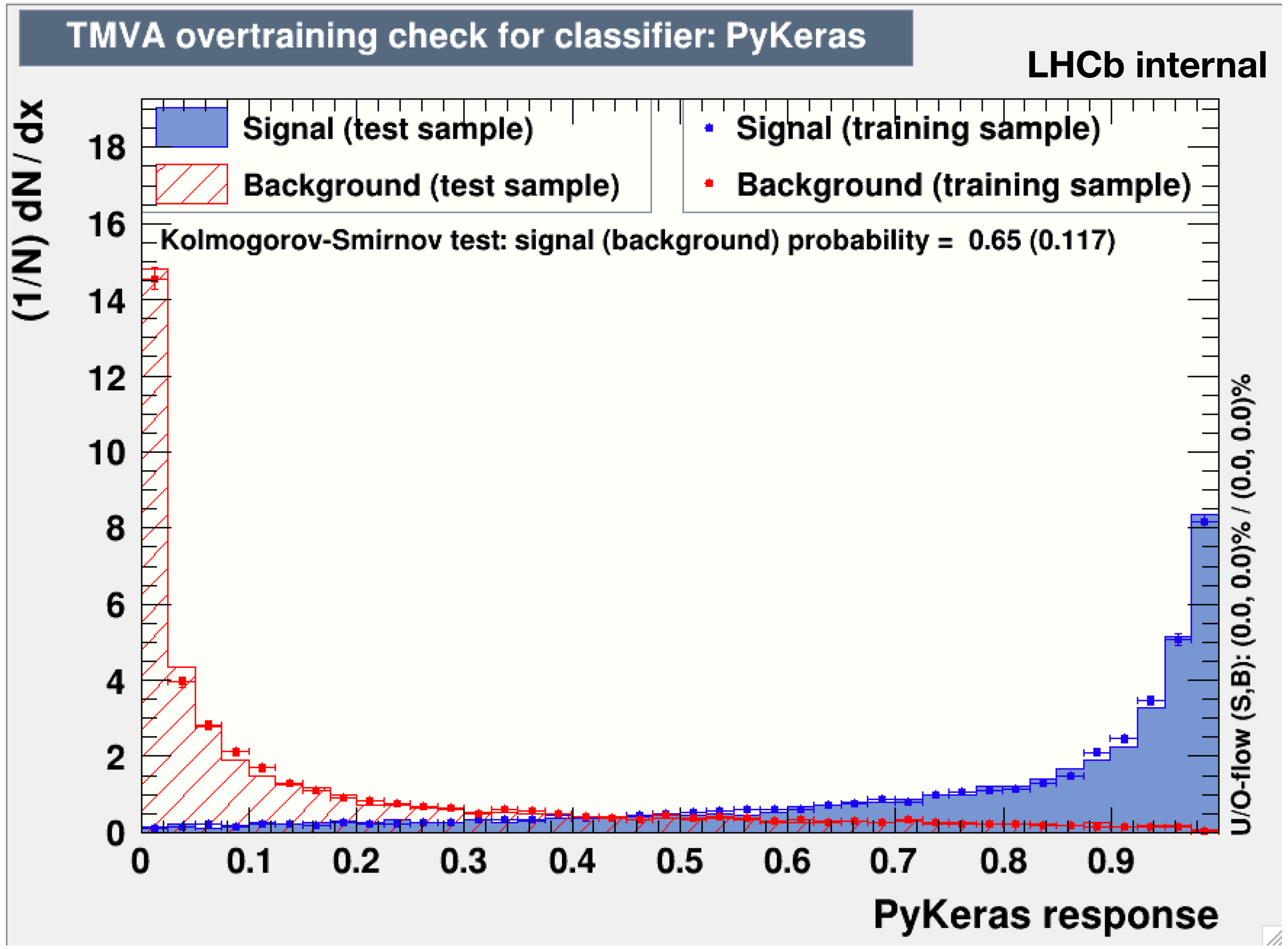
- Invariant mass of signal candidate plus highest ranked track.
- Invariant mass of signal candidate plus the two highest ranked tracks.
- MVA value of highest ranked track.
- . . .

DNN performance

Compare different MVA methods



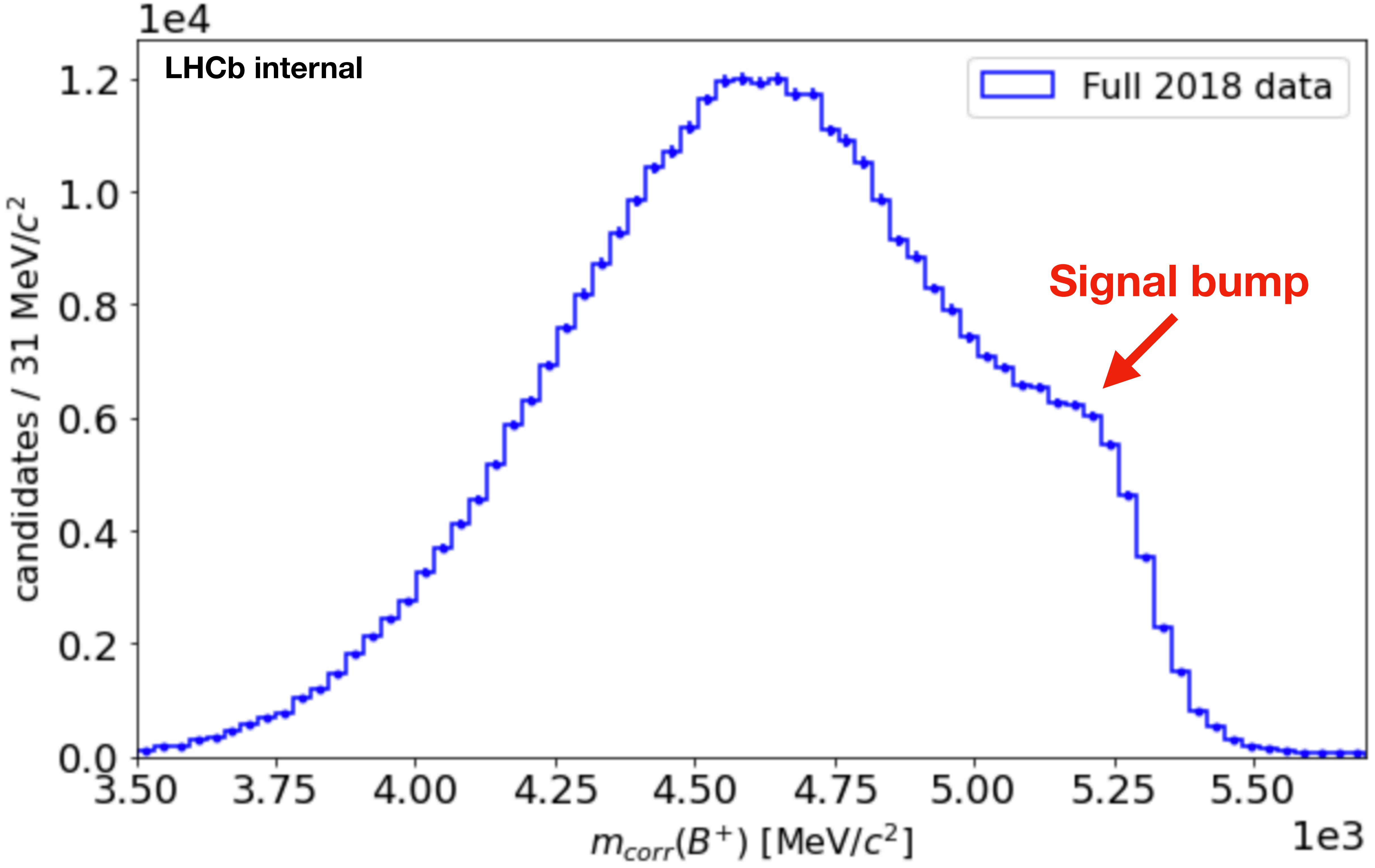
DNN output variable

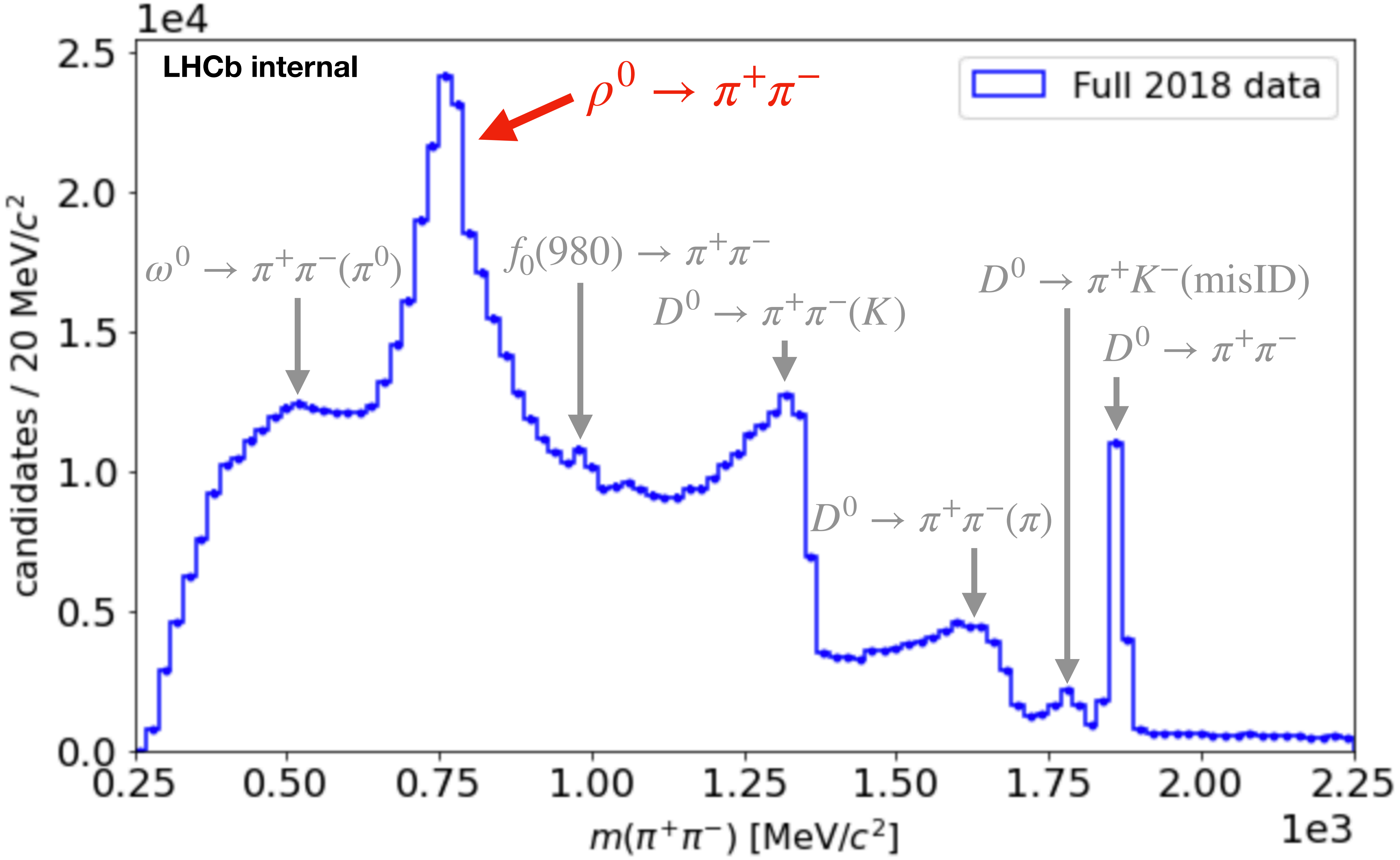


→ Good separation power between signal and background.

After DNN selection

- **Distributions after applying (non-optimised) selection cut:**





- **Backgrounds left after DNN selection are simulated and included in the signal fit.**

Normalisation channel fit

Reminder:

- We want to measure the differential decay rate of $B^+ \to \rho^0[\to \pi^+\pi^-]\mu^+\nu_\mu$ where we use $B^+ \to D^0[\to \pi^+\pi^-]\mu^+\nu_\mu$ as a normalisation channel.

Template fit, we extract the templates from:

1.) **Simulation of normalisation channel** :

$B^+ \to \bar{D}^0\mu^+\nu_\mu$ with $D^0 \to \pi^+\pi^-$

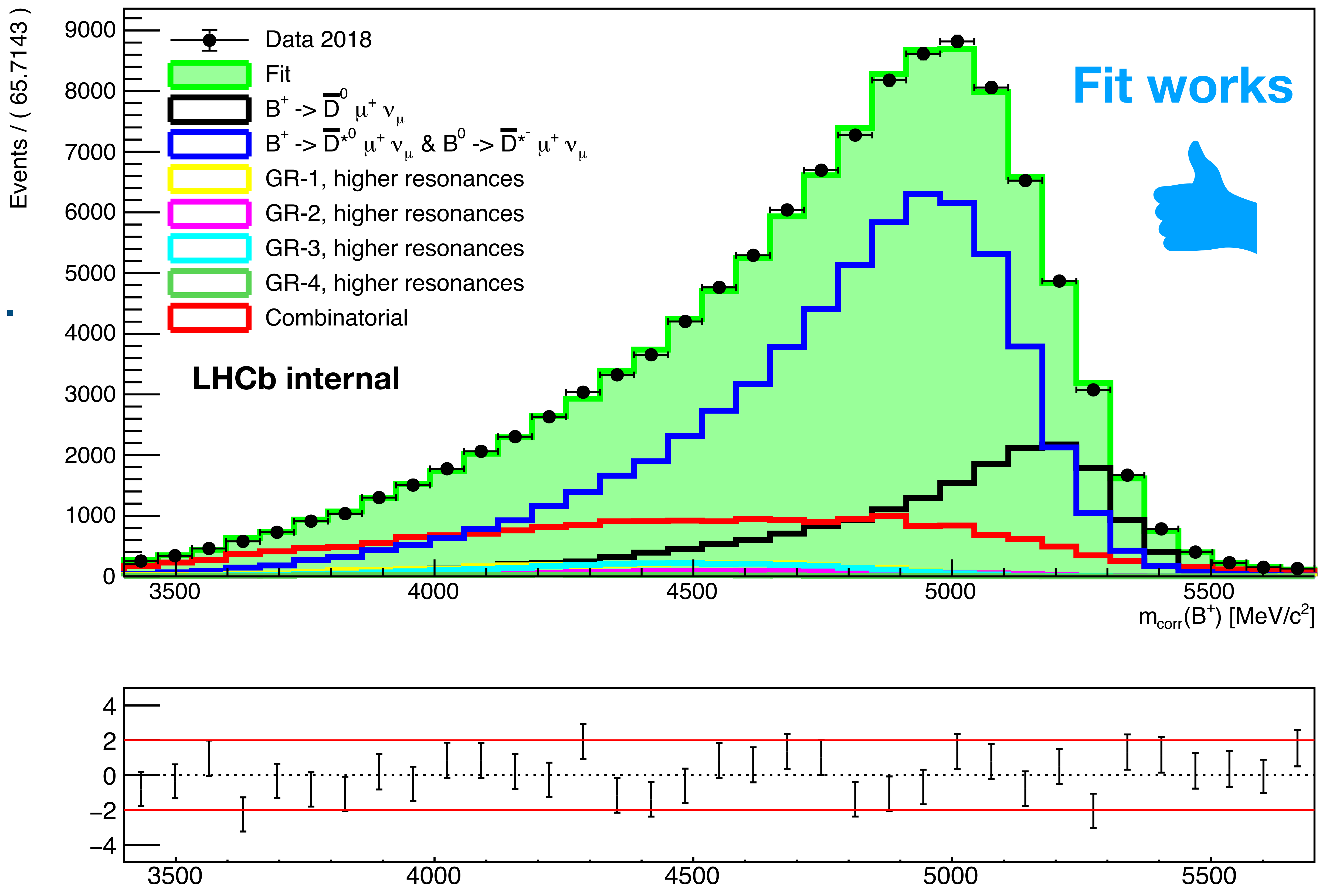
2.) **Simulations of backgrounds** :

$B \to \bar{D}^{(*,**)}\mu^+\nu_\mu X$ modes.

3.) **Combinatorial background from data** :

side-bands of the $m(D^0)$

- We use **all templates** to fit **the full 2018 data** sample.



Fit works

Fit result :

$S = 19586 \pm 434$ (~2.2% relative error)

Preliminary fit of signal

Template fit, we extract all templates from simulations of:

$B^+ \to \rho^0 \mu^+ \nu_\mu$ (signal)

$B^0 \to J/\psi(\to \mu^+ \mu^-)\rho^0$ (control channel)

- **Semileptonic V_{ub} backgrounds:**

  $B^+ \to \omega^0 \mu^+ \nu_\mu$

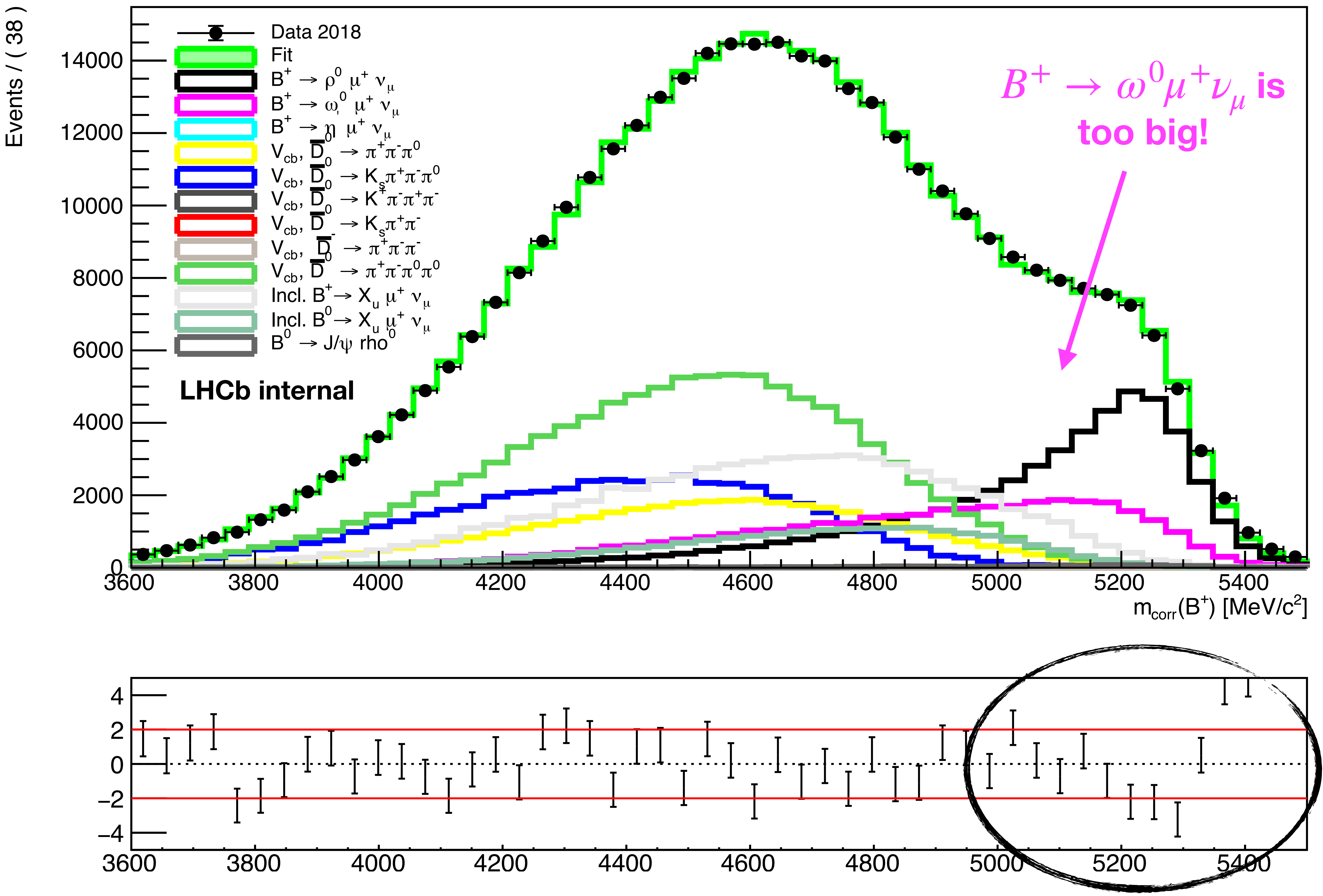
  $B^+ \to \eta^{'} \mu^+ \nu_\mu$

  Inclusive V_{ub} samples

- **Semileptonic V_{cb} decays where:**

  $D \to \pi^+ \pi^- X$ with X being 1-2 charged or neutral particles.

- We use **all templates** to fit **the full 2018 data** sample.



$B^+ \to \omega^0 \mu^+ \nu_\mu$ is too big!

Room for improvements..

There seems to be a missing component at high $m_{corr}(B^+)$

Conclusion and outlook

Working towards measuring the $B^+ \to \rho^0 \mu^+ \nu_\mu$ differential decay rate and extract $|V_{\mathrm{ub}}|$.

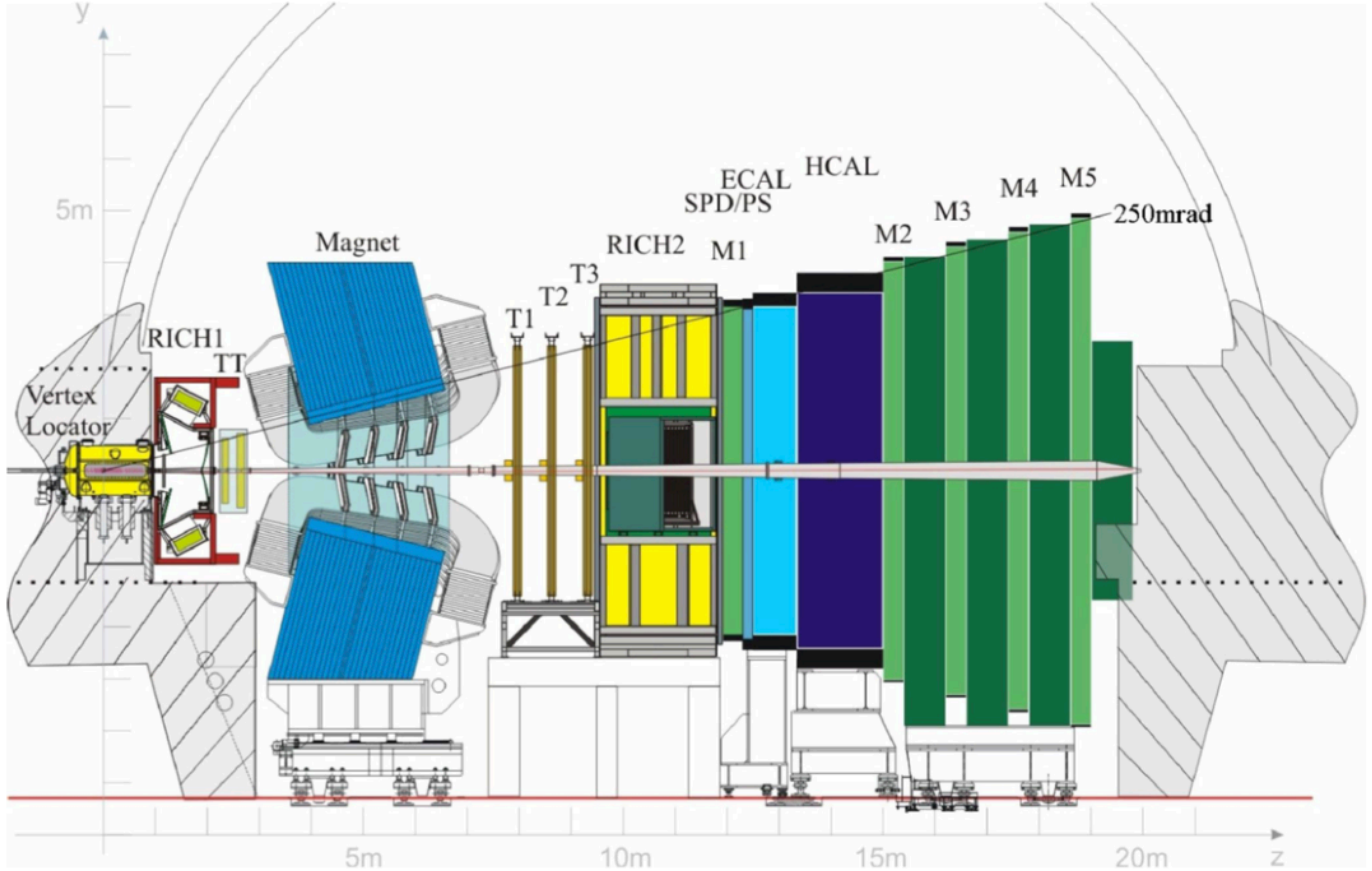
- Such a measurement can help us understand the $|V_{\mathrm{ub}}|$ **puzzle** and a possible **new physics** explanation.
- We have developed a **DNN** that **efficiently isolates signal against the dominant background** of $b \to c$ semileptonic decays.
- We have **successfully fitted our normalisation channel,** and **performed a preliminary fit of signal that looks promising**.
- Next, perform fit in **bins of** q^2, compute **differential decay rate** and **extract** $|V_{\mathrm{ub}}|$.

Thank you for your attention :)

Back-up slides

Experimental realisation: The LHCb experiment

- **Measure the $B^+ \to \rho^0(\to \pi^+\pi^-)\mu^+\nu_\mu$ differential decay rate using Run 2 data collected by the LHCb experiment**



The LHCb detector in Run 2 configuration [?].

[1] LHCb collaboration, Int. J. Mod. Phys. A 30, 1530022 (2015).

NP: right-handed weak current and observables

Lagrangian with right-handed weak current

- Effective Lagrangian allowing for a right handed admixture to the SM weak current [2]:

$$\mathcal{L}_{\rm eff} = -\frac{4G_F}{\sqrt{2}} V_{ub}^L \big(\bar{u}\gamma_\mu P_L b + \epsilon_R\, \bar{u}\gamma_\mu P_R b\big)\big(\bar{\nu}\gamma^\mu P_L \ell\big) + \text{h.c.},$$

Decay	$\lvert V_{ub}\rvert \times 10^3$	ϵ_R dependence
$B \to \pi\, \ell\bar{\nu}$	3.23 ± 0.30	$1 + \epsilon_R$
$B \to X_u \ell\bar{\nu}$	4.39 ± 0.21	$\sqrt{1+\epsilon_R^2}$
$B \to \tau\, \bar{\nu}_\tau$	4.32 ± 0.42	$1 - \epsilon_R$
Decay	$\mathcal{B} \times 10^4$	
$B \to \rho\, \ell\bar{\nu}$	$1.97 \pm 0.16\ (q^2 < 12\,\text{GeV}^2)$	
$B \to \omega\, \ell\bar{\nu}$	$0.61 \pm 0.11\ (q^2 < 12\,\text{GeV}^2)$	

[2] Florian U. Bernlochner et al. Phys. Rev. D 90, 094003 (2014).

Examples of simpler observables

These amounts to counting experiments where the partial branching fraction is determined in different regions of phase space and this information is then combined to:

(i) construct asymmetries sensitive to NP, e.g. :

$$S = \frac{A - B}{A + B}.$$

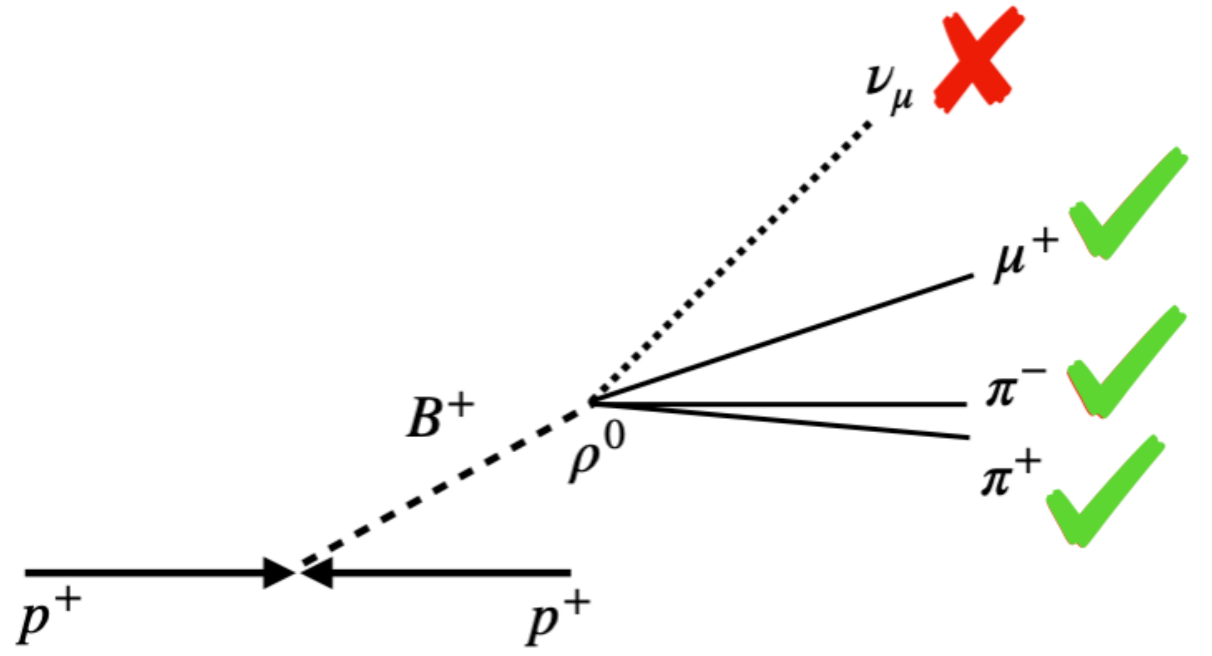
A and B are the decay rates in two different regions of $\{\cos\theta_l, \cos\theta_V\}$ phase space (integrated over ϕ and q^2).

(ii) isolate the J_i coefficients and construct ratios, e.g. :

$$J_i = \frac{1}{N_i} \sum_{j=1}^{8} \sum_{k,l=1}^{4} \eta^{\chi}_{i,j}\, \eta^{\theta_\ell}_{i,k}\, \eta^{\theta_V}_{i,l} \left[\chi^{(j)} \otimes \theta_\ell^{(j)} \otimes \theta_V^{(k)}\right], \quad \langle P_{i,j}\rangle_{\rm bin} = \frac{\int_{\Delta q^2} dq^2 J_i}{\int_{\Delta q^2} dq^2 J_j}.$$

[2] Florian U. Bernlochner et al. Phys. Rev. D 90, 094003 (2014).

Corrected B mass



- We cannot reconstruct the invariant mass of the B meson due to the **unmeasured neutrino**.

- We can reconstruct the so-called visible mass corresponding to the invariant mass of the visible final state particles, the muon and the rho meson, however, this is not a good discriminating variable, since the distribution is very broad.

- To **compensate for the unmeasured neutrino** we can apply a **kinematic correction** to the visible mass and obtain the so-called corrected B mass, which is a better discriminating variable due to its **narrower distribution**.:

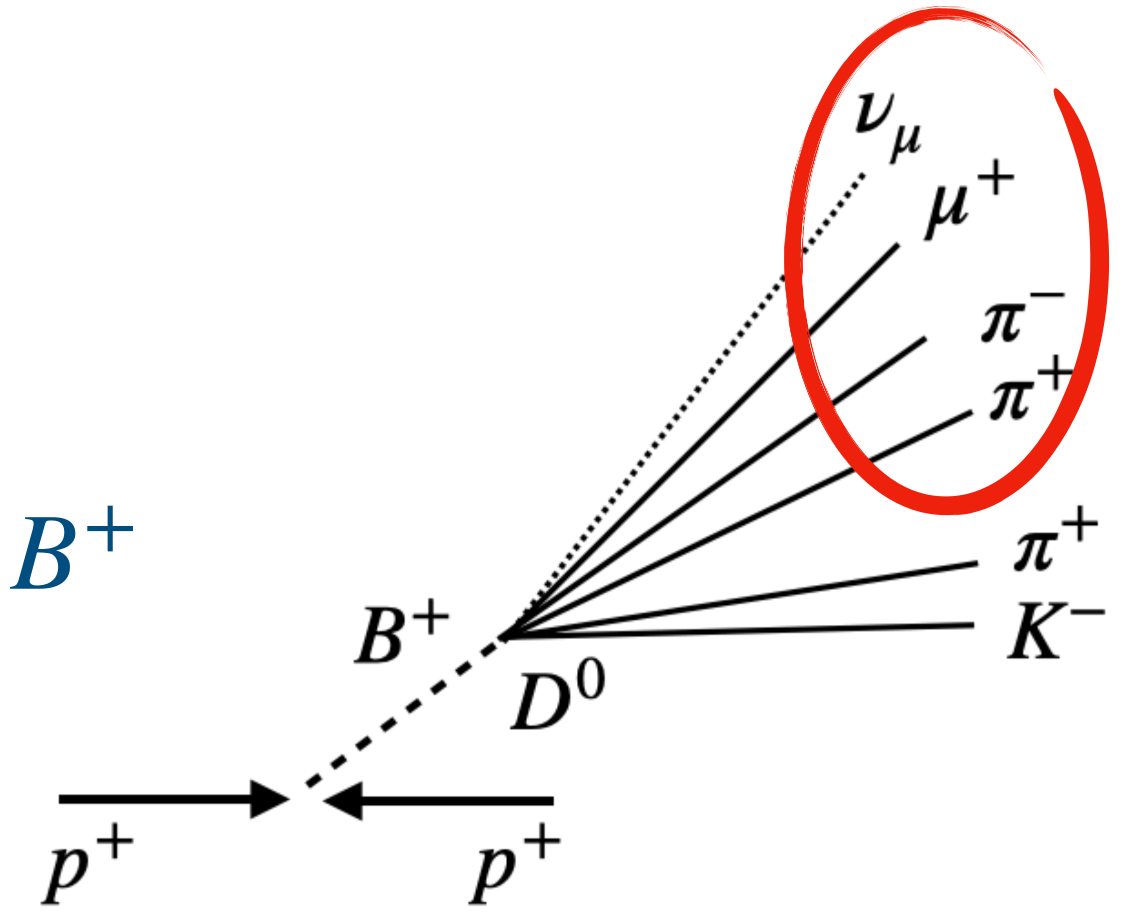
- $m_{corr}(B^+) = \sqrt{m_{vis}^2 + p_\perp^2} + p_\perp$

- Here $p_\perp$ is the momentum of the final state particle perpendicular to the flight direction of the B+ meson, and with momentum conservation it is equal to the momentum of the neutrino perpendicular to the flight direction of the B+ meson.

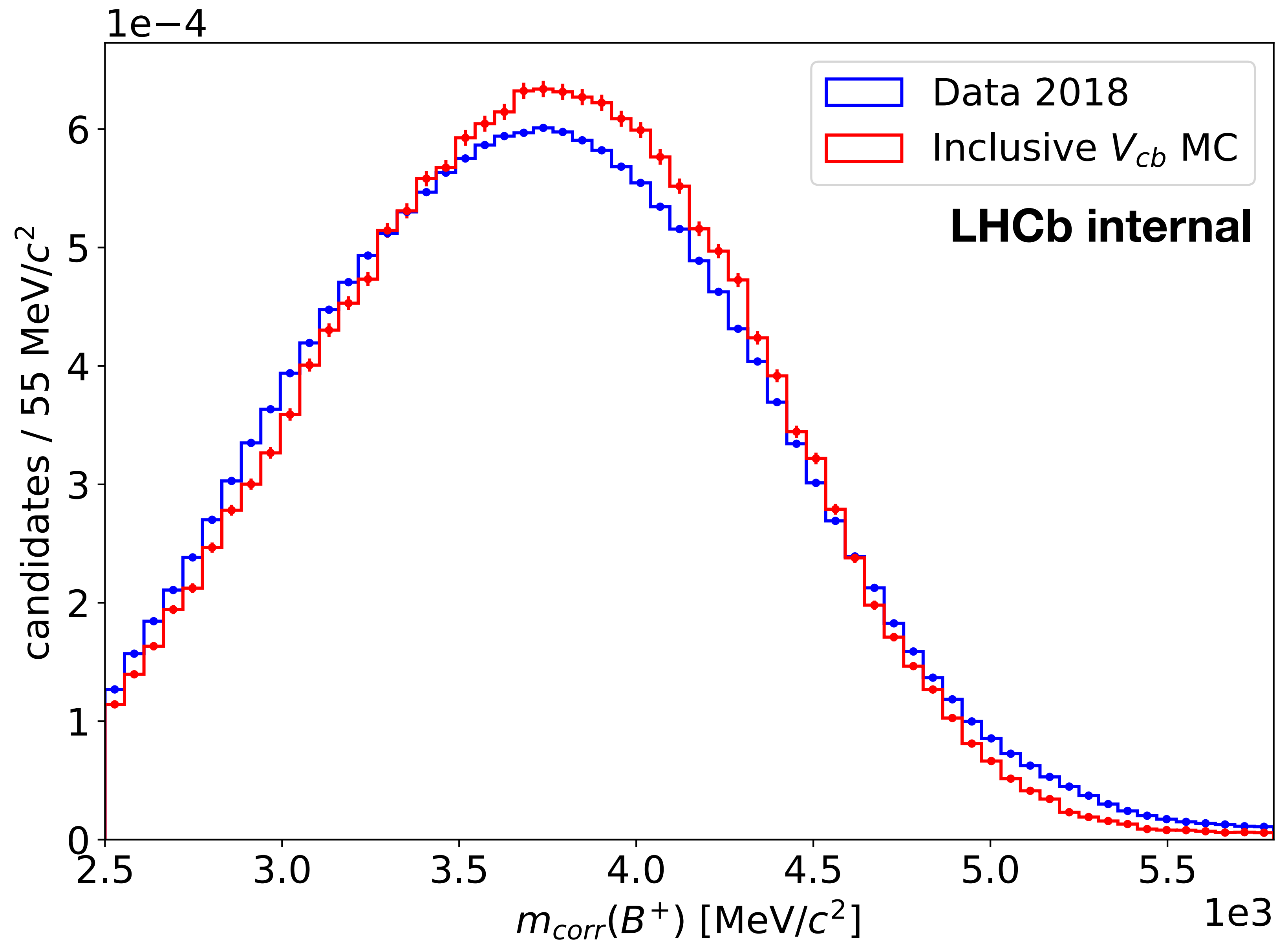
Variable introduction: https://arxiv.org/pdf/hep-ex/9708015.pdf

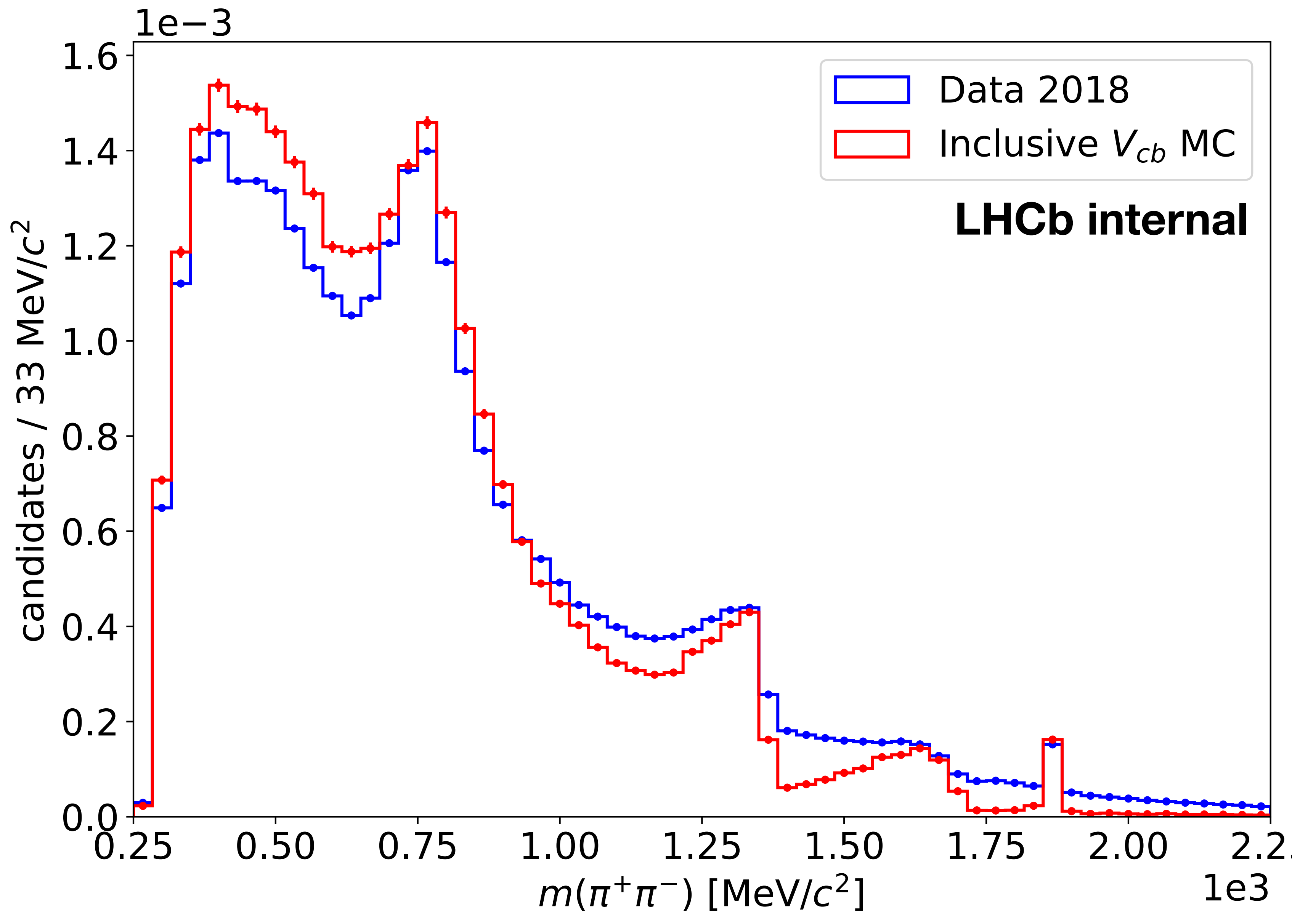
Derivation: https://lphe.epfl.ch/publications/theses/Lino_FerreiraLopes_MasterProject.pdf

Inclusive V_{cb} simulation



- **Describe dominant background:** inclusive V_{cb} simulation consisting of simulated B^+ decays into $\mu^+\nu_\mu(X)$ and either D^0, D^{*0} or D^{**0} with a $\pi^+\pi^-X$ final state.





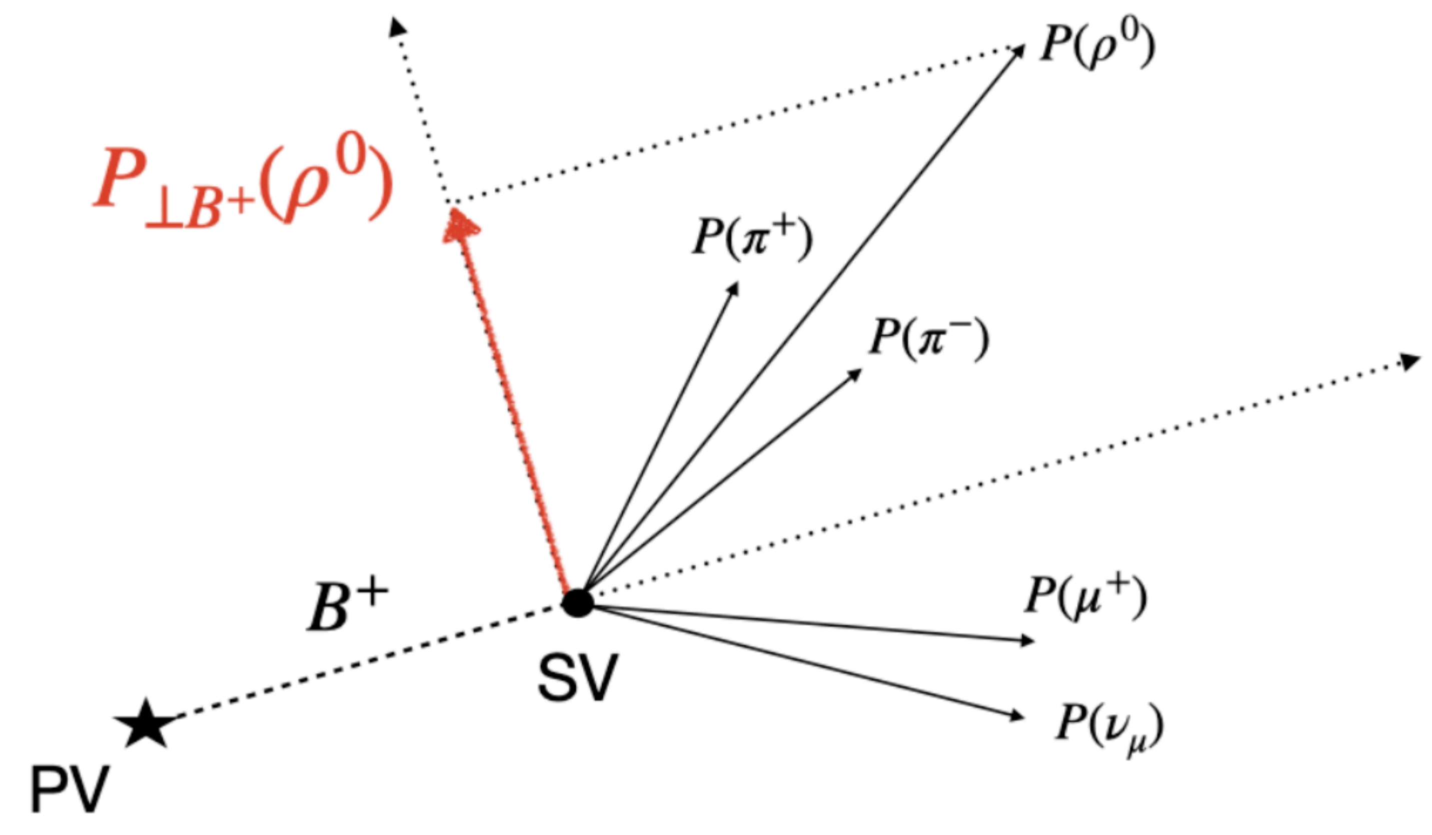
- **Inclusive V_{cb} simulation** and data has similar $m_{corr}(B^+)$ and $m(\pi^+\pi^-)$ distributions.

"Good background sample for MVA selection"

Diagonal cuts: explanation and motivation

1. Diagonal cut

- We impose a requirement on the component of the ρ^0 momentum transverse to the B^+ flight direction, $P_{\perp B^+}(\rho^0)$, as a function of $m_{\rm corr}(B^+)$ [4].



- This cut removes a large fraction of the background while retaining a high signal efficiency.

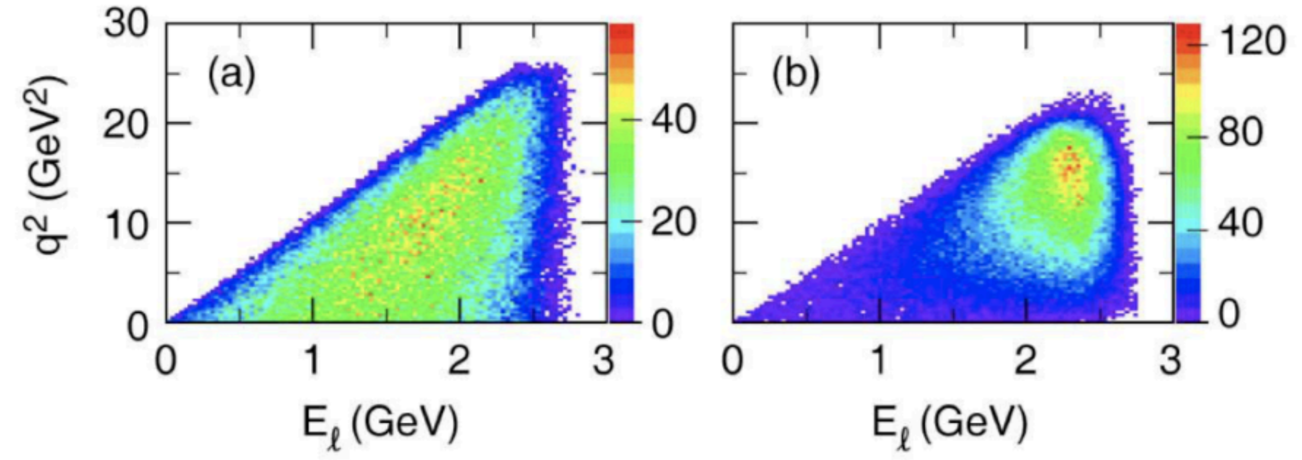
[4] T. J. Boettcher et al., LHCb-ANA-2016-068, (2017).

2. Diagonal cut

- We impose a requirement on the momentum of the μ^+ in the B^+ rest frame, $P(\mu_{B^+ rest})$, as a function of $m_{\rm corr}(B^+)$.

- In a paper from BaBar [5], they show that the 2D distribution of q^2 and the charge lepton energy in the B rest frame is different for $B \to \rho l \nu$ (Fig. b) and $B \to \pi l \nu$ (Fig. a) due to their different spin structure.

$$\frac{d\Gamma(B\to\rho\ell\nu)}{dq^2\, d\cos\theta_{W\ell}} = |V_{ub}|^2 \frac{G_F^2\, p_\rho\, q^2}{128\pi^3 M_B^2} \times \left[\sin\theta_{W\ell}^2 |H_0|^2 + (1-\cos\theta_{W\ell})^2 \frac{|H_+|^2}{2} + (1+\cos\theta_{W\ell})^2 \frac{|H_-|^2}{2}\right].$$



- The $B\bar{B}$ rest frame cannot be exactly recovered in hadron colliders, so we approximate the B rest frame as in ANA-note $R(D^*)$ muonic. [6].

- We assume that the proper velocity $\gamma\beta$ of the visible part of the decay (Y) along the z-axis is equal to the one of the B meson, and we get: $(P_B)_z = \frac{m_B}{m_Y}(P_Y)_z$

- Using the unit vector between the primary vertex and B decay vertex, we get: $|P_B| = \frac{m_B}{m_Y}(P_Y)_z\sqrt{1+\tan^2\alpha}$, which is then used to boost the μ back in the B rest frame.

[5] The BABAR Collaboration, arXiv:1005.3288 [hep-ex], (2010).
[6] B. Hamilton et al., LHCb-ANA-2014-0527, (2015).

Diagonal cuts

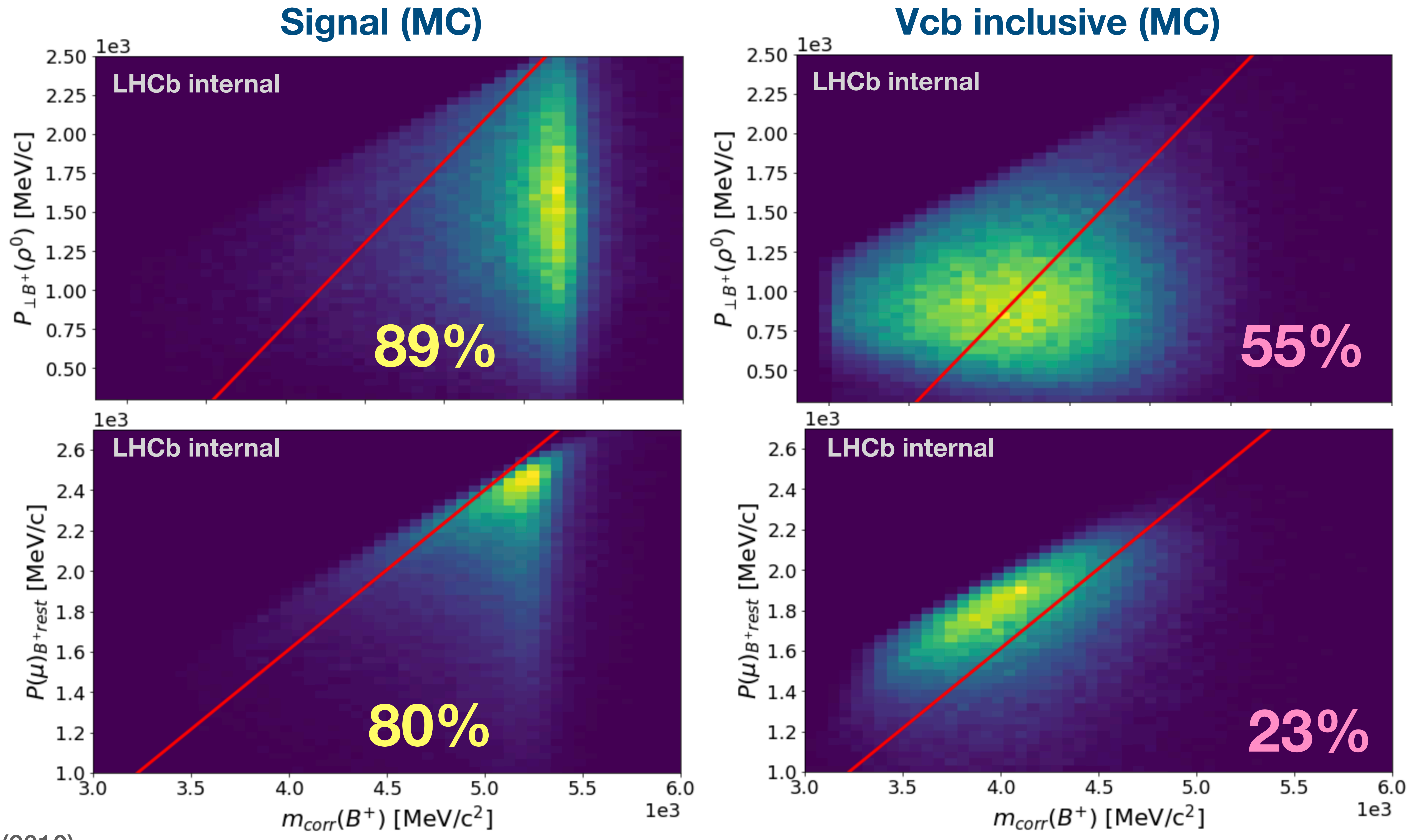
- **Goal:** reduce ratio of background-to-signal, also in the low $m_{corr}(B^+)$ region.

1.) Diagonal cut:

$P_{\perp B^+}(\rho^0)$ versus $m_{corr}(B^+)$ [1]

2.) Diagonal cut:

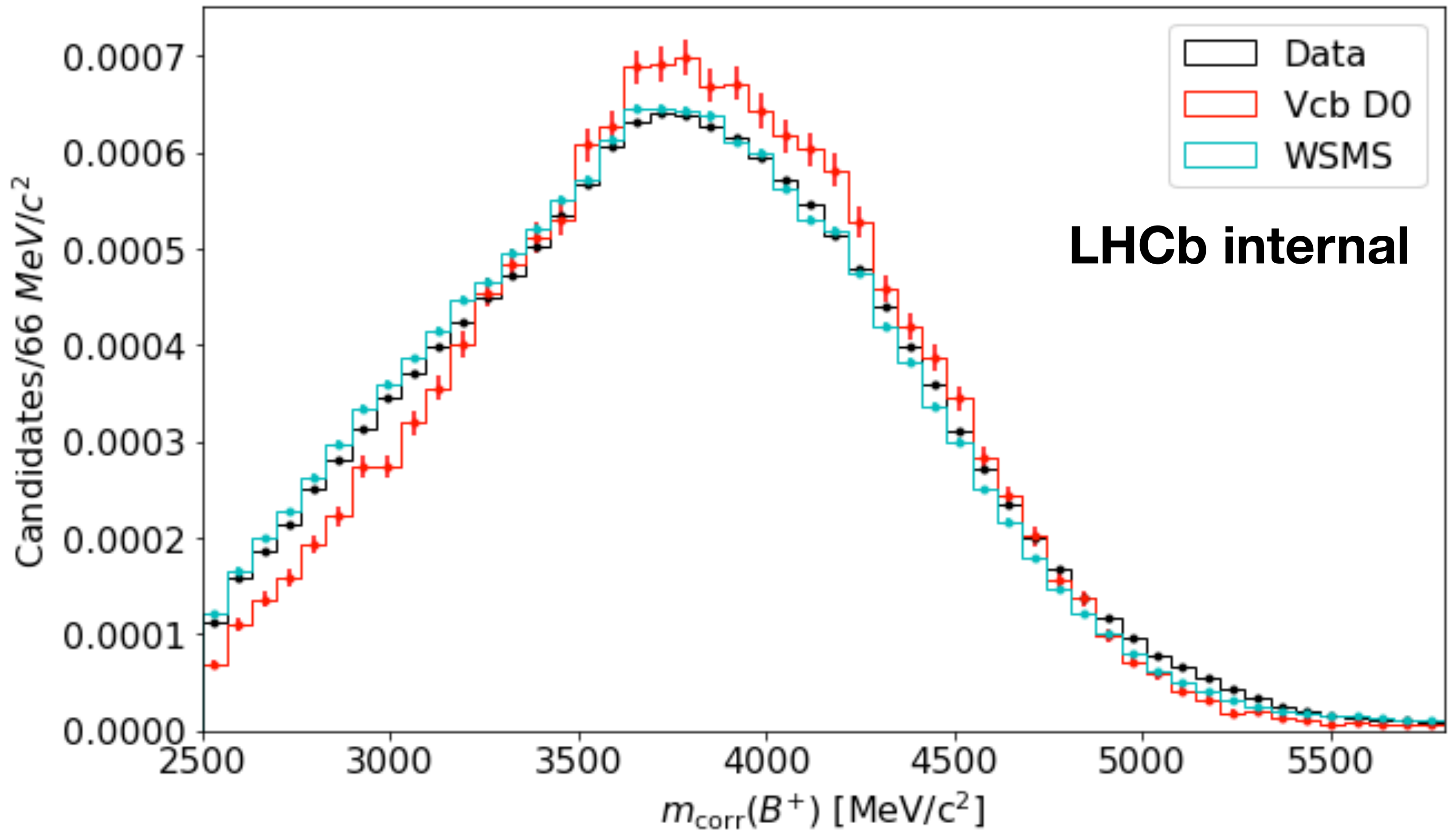
$P_{B^+rest}(\mu^+)$ versus $m_{corr}(B^+)$ [2],[3]

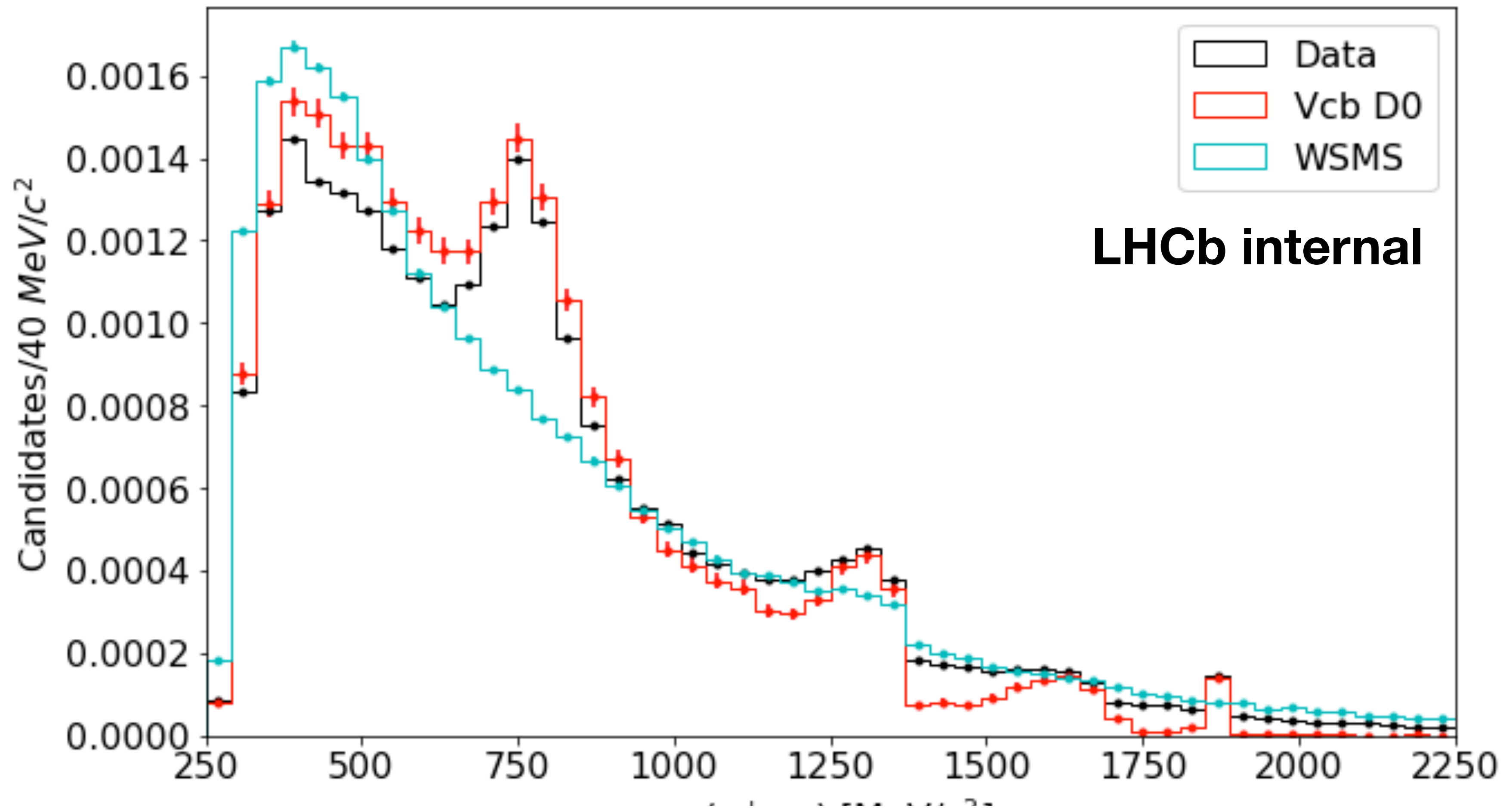


[1] T. J. Boettcher et al., LHCb-ANA-2016-068, (2017).
[2] The BABAR Collaboration, arXiv:1005.3288 [hep-ex], (2010).
[3] B. Hamilton et al., LHCb-ANA-2014-0527, (2015).

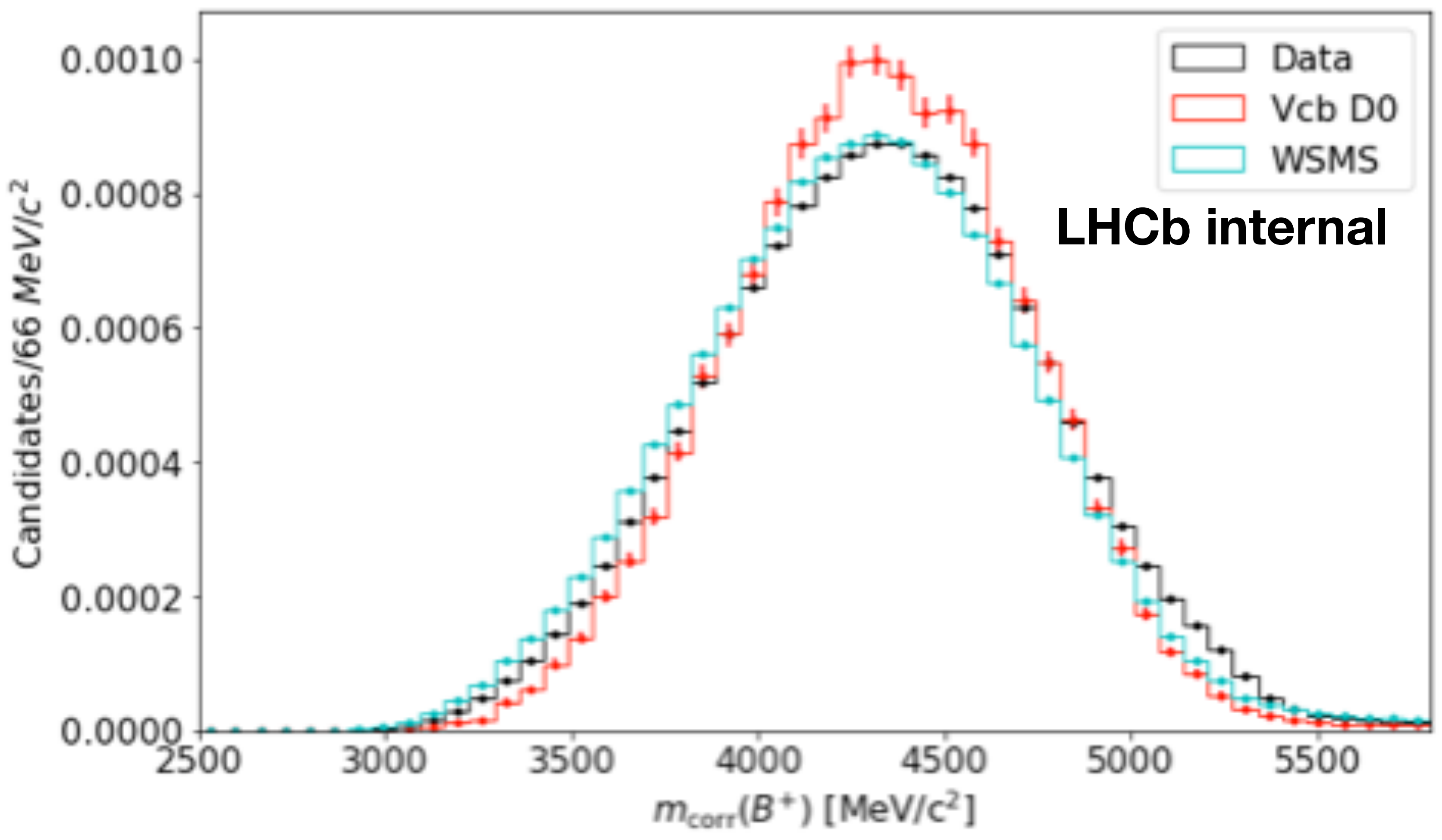
Effect of diagonal cuts

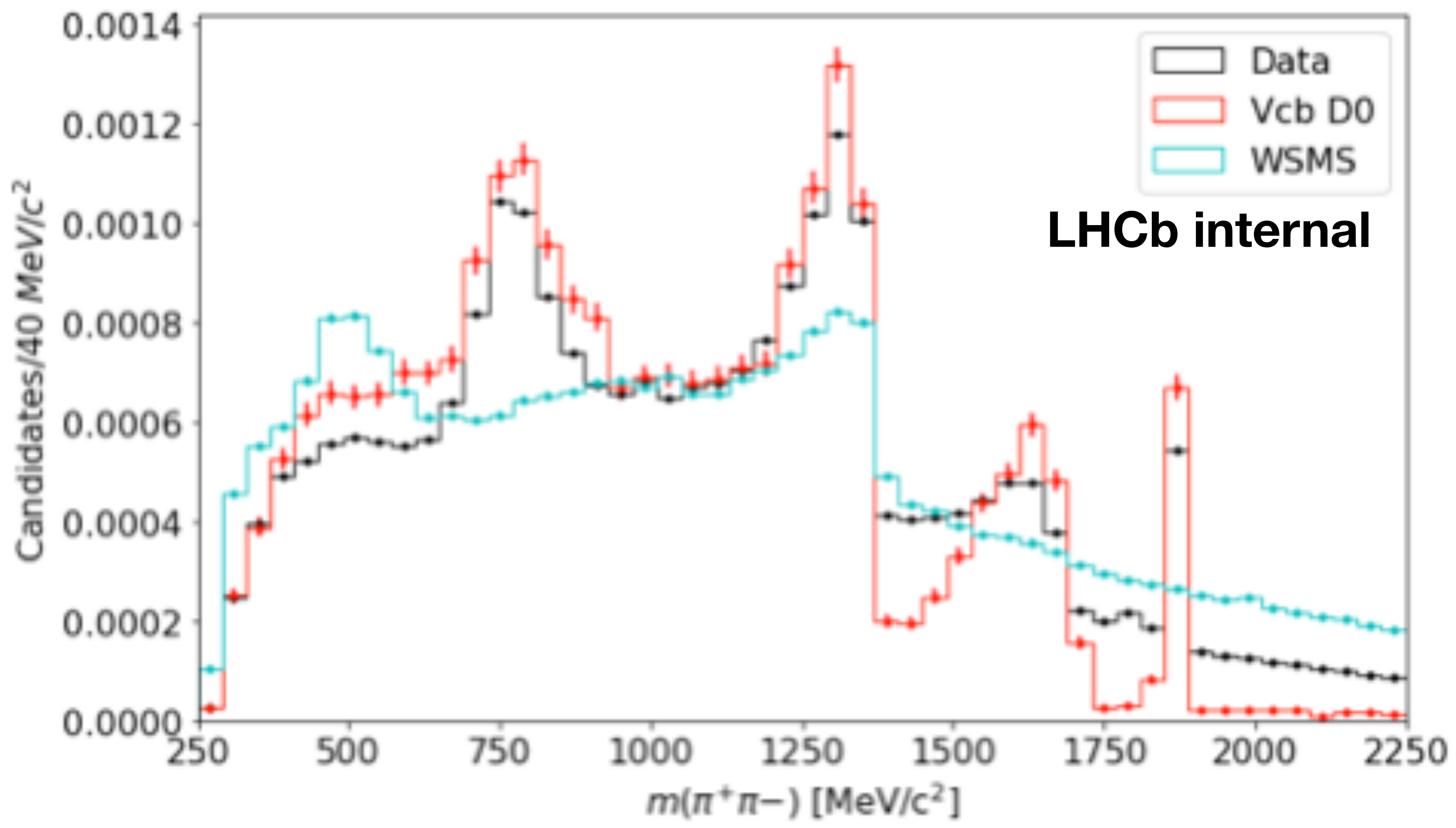
Before :





After :





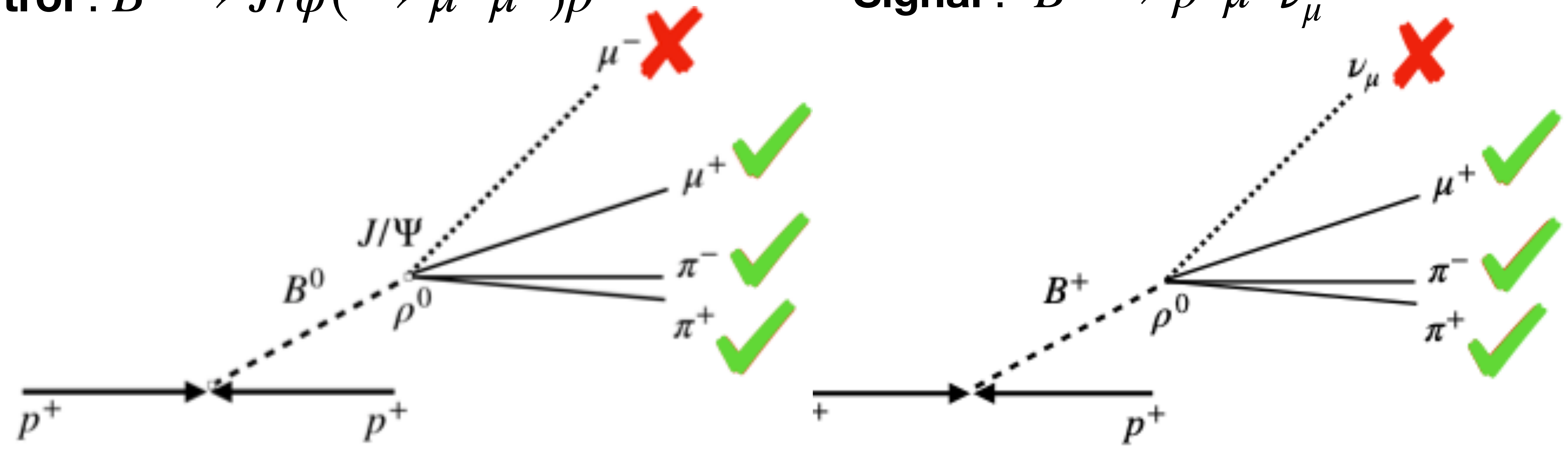
- After preselection Vcb-D0 and data still have very similar $m_{corr}(B^+)$ and $m(\pi^+\pi^-)$ distributions

"Good background sample for MVA selection"

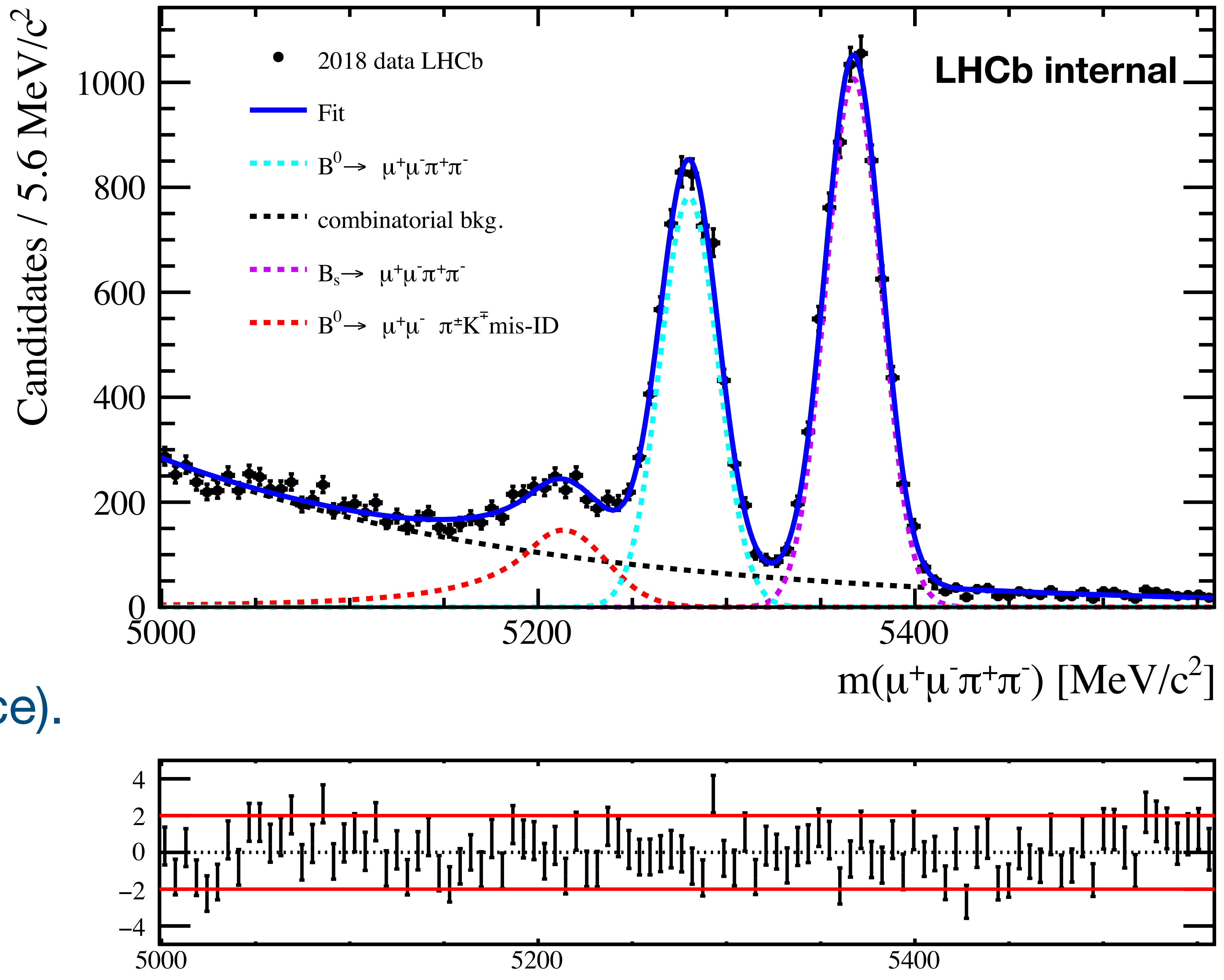
Control channel

- **Control channel:** $B^0 \to J/\psi(\to \mu^+\mu^-)\rho^0$ is used to verify signal MC for MVA selection.
- With **one muon missing**, the control channel has the **same topology** and **same visible final state** as signal.

Control : $B^0 \to J/\psi(\to \mu^+\mu^-)\rho^0$ **Signal** : $B^+ \to \rho^0\mu^+\nu_\mu$



The fit is good, and we obtain the s-weights



- The **main difference** between the two modes is the physics of the **leptonic system**: W-boson (weak force) vs. γ (electromagnetic force).
- Control channel can be **fully reconstructed in data** by adding the most compatible muon to the signal candidate.
- After applying weights to correct for known data-MC differences, we find **MC to be consistent with data**.

Building and training DNN

Method:
Deep neural network (TMVA/keras)

Architecture:
input layer (8 variables)
1. Hidden layer (100 neurones)
2. Hidden layer (50 neurones)
output layer (2, softmax)

Variable transformation :
G,D,G,D (G: gaussian, D: decorrelation)

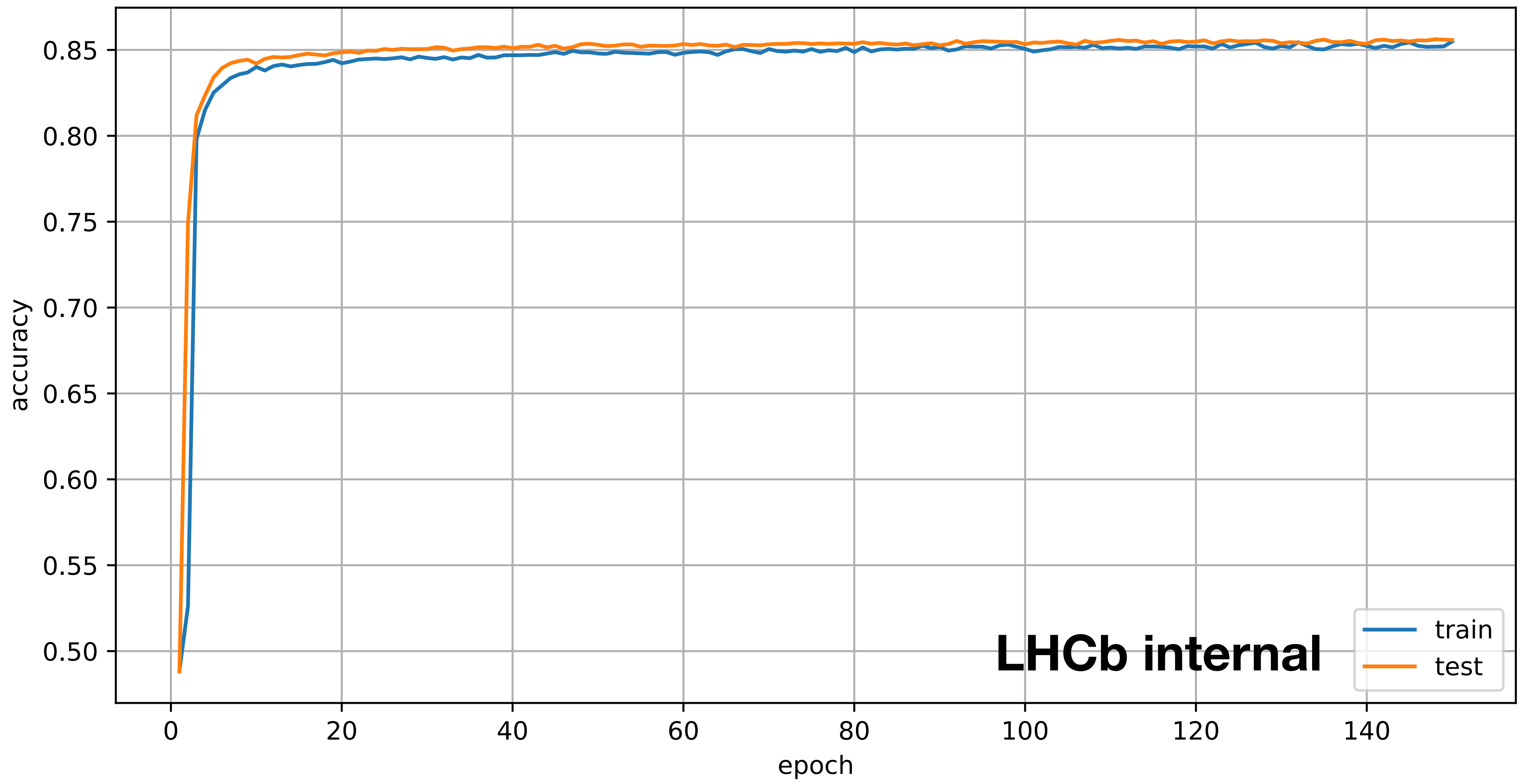
Train/test:
Split sample: 50% / 50%.

Epochs : 250

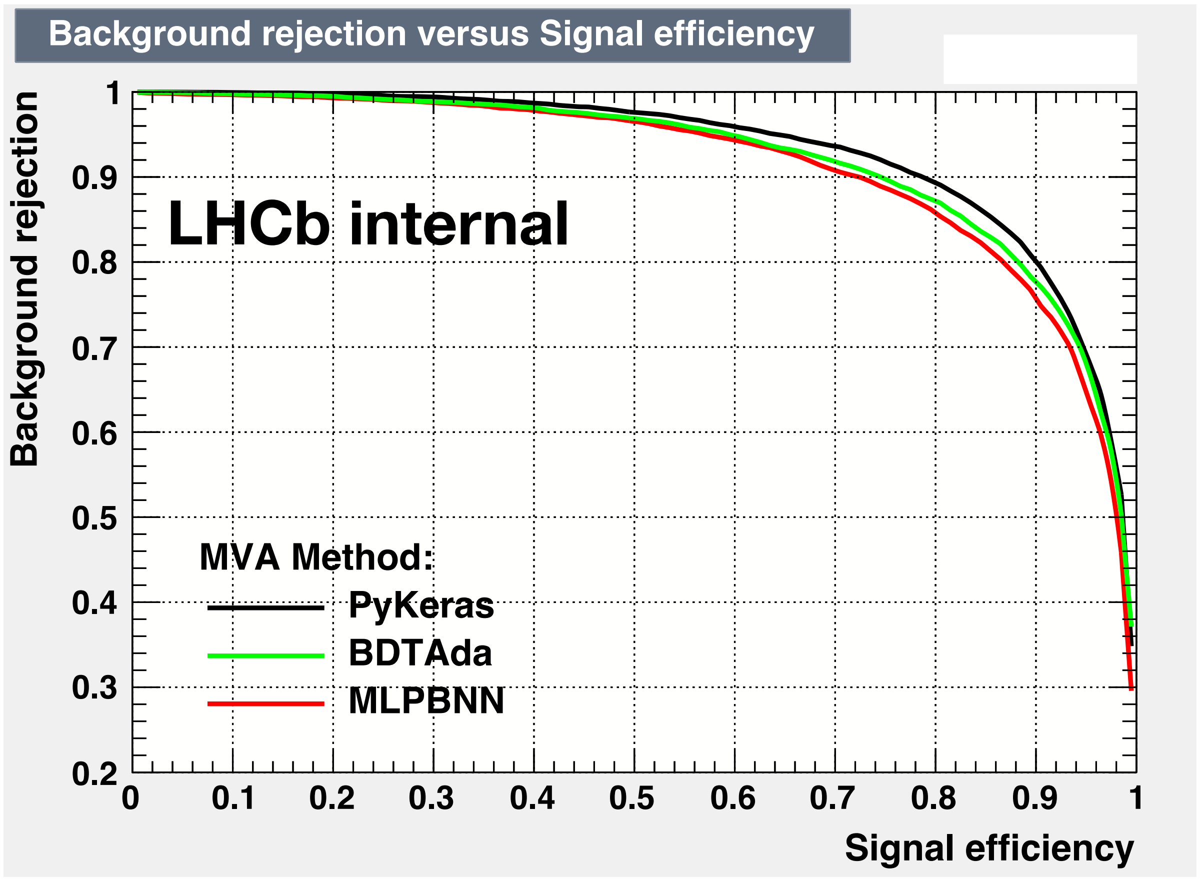
Batch Size : 32

Evaluering: AUC = 0.94

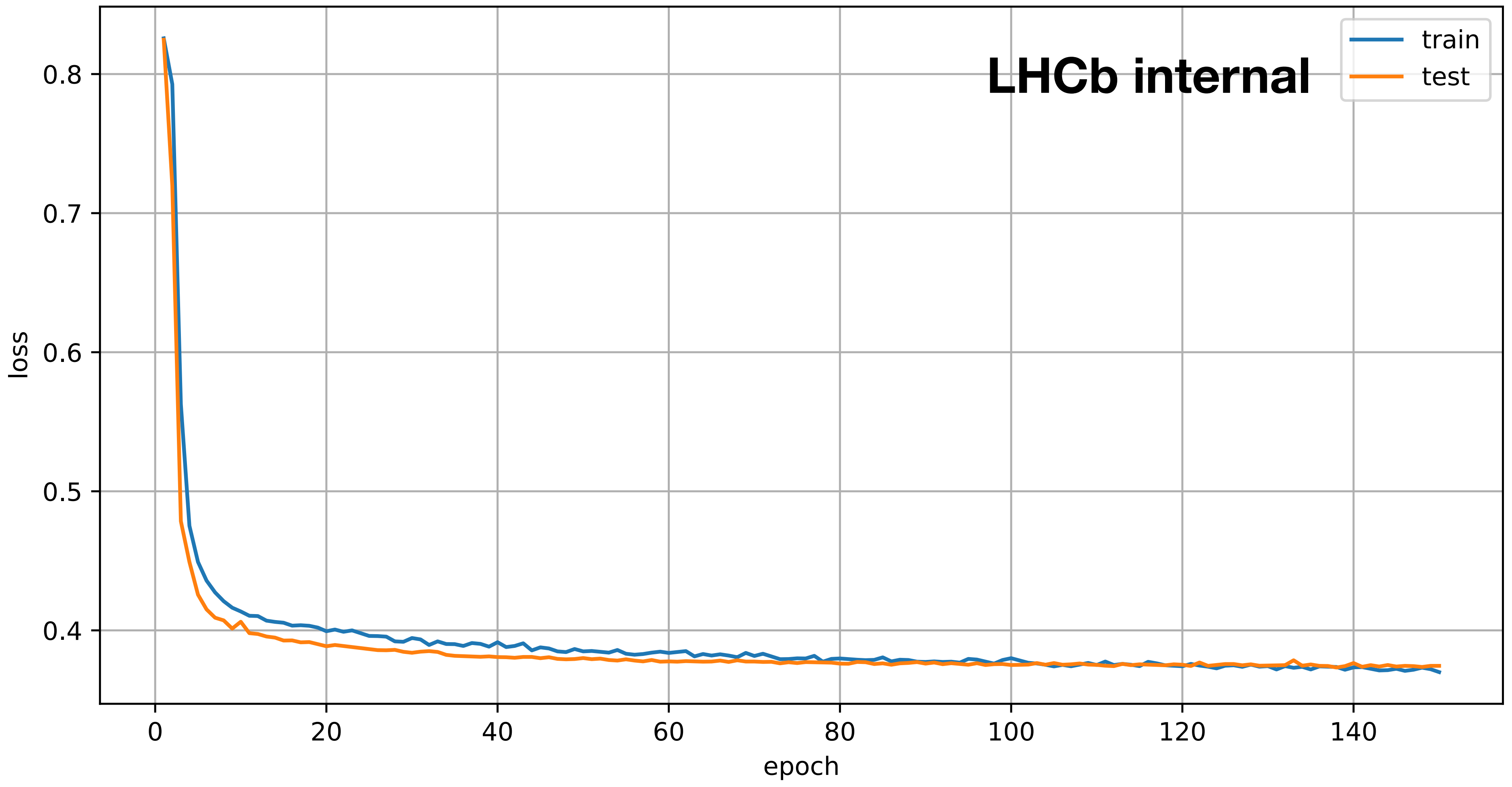
Accuracy curves



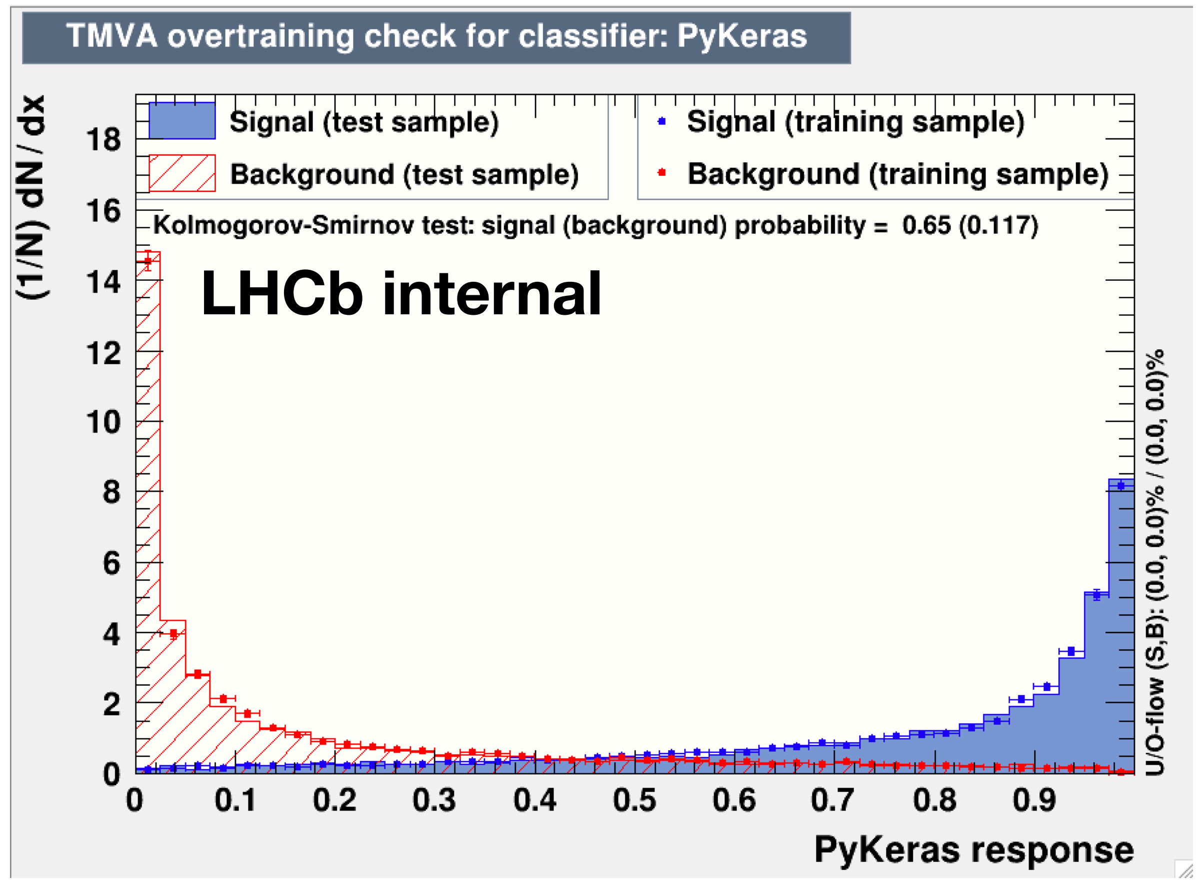
ROC curves



Loss curves



DNN output variable



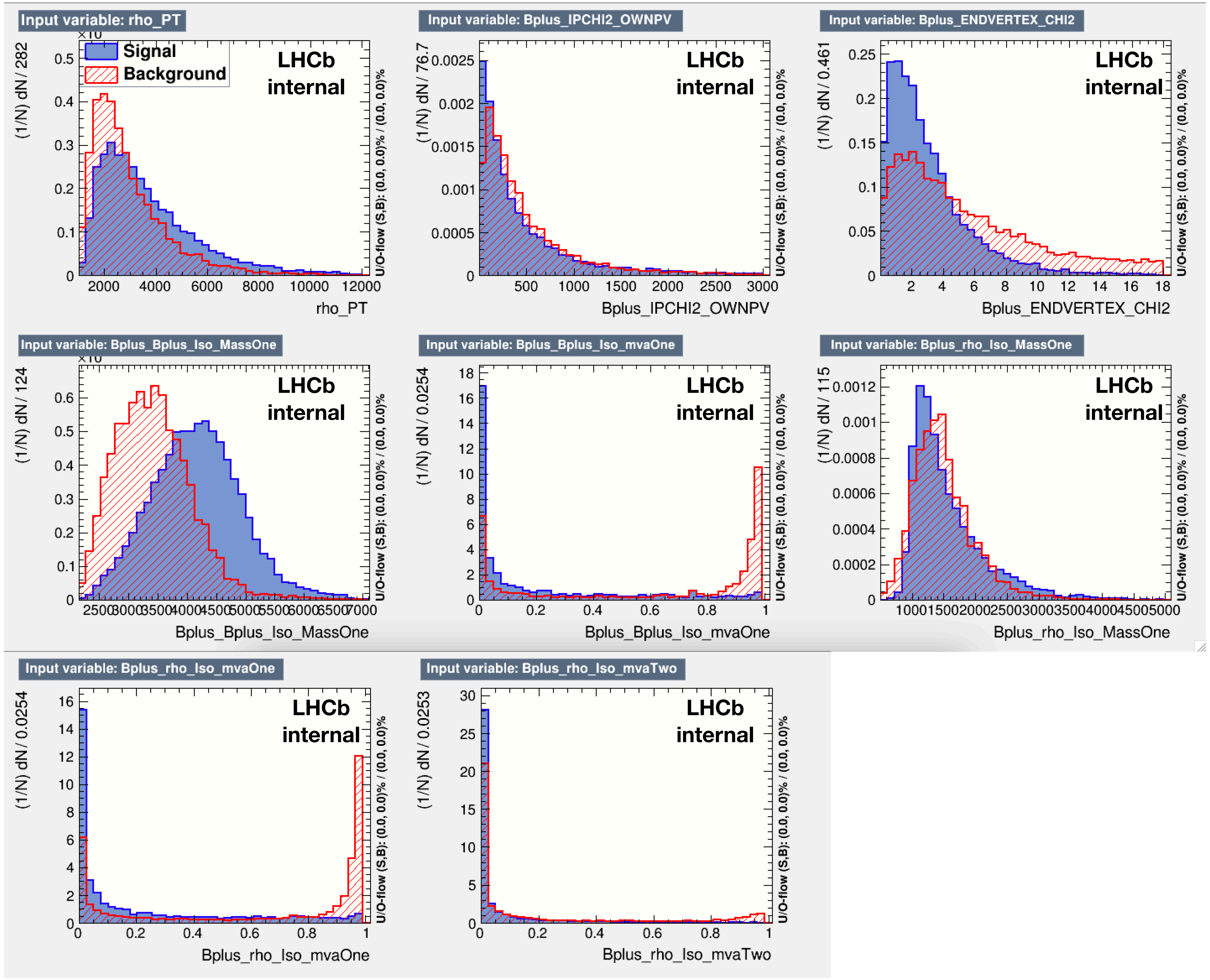
DNN variable distributions

Input samples:

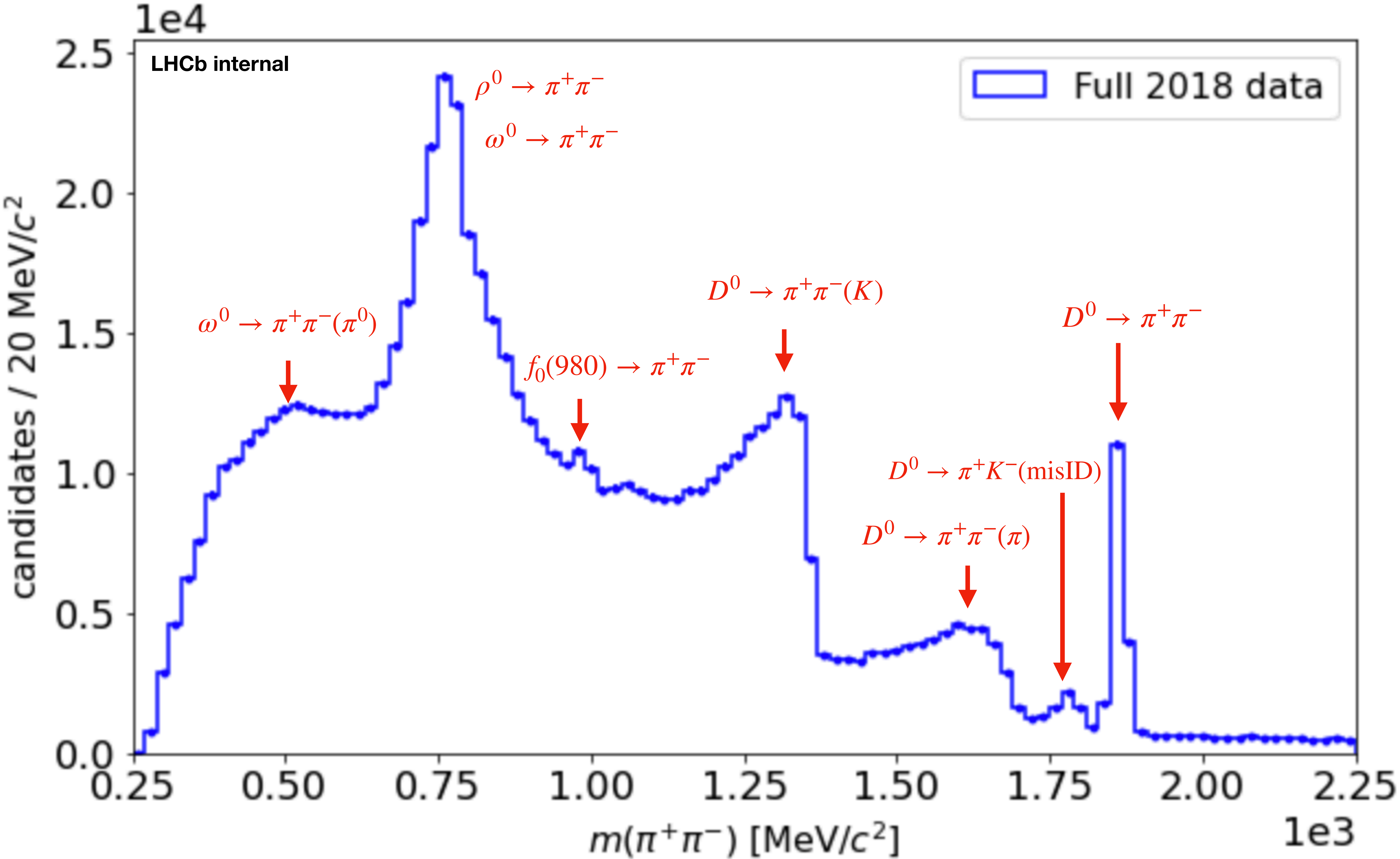
- **Signal:** $B^+ \to \rho^0 \mu^+ \nu_\mu$ MC (~ 22 k)
- **Background:** inclusive Vcb MC (~ 22 k)

Input variables:

- $P_T(\rho^0)$
- $IP\chi^2(B^+)$, IP: impact parameter.
- $EV\chi^2(B^+)$, EV: end-vertex.
- $Mass(B^+ + \text{track 1})$, M: invariant mass.
- $mvaVal(B^+ + \text{track 1})$, mva-ranking value.
- $Mass(\rho^0 + \text{track 1})$
- $mvaVal(\rho^0 + \text{track 1})$
- $mvaVal(\rho^0 + \text{track 1 \& 2})$



$\pi^+\pi^-$-mass distribution after DNN cut



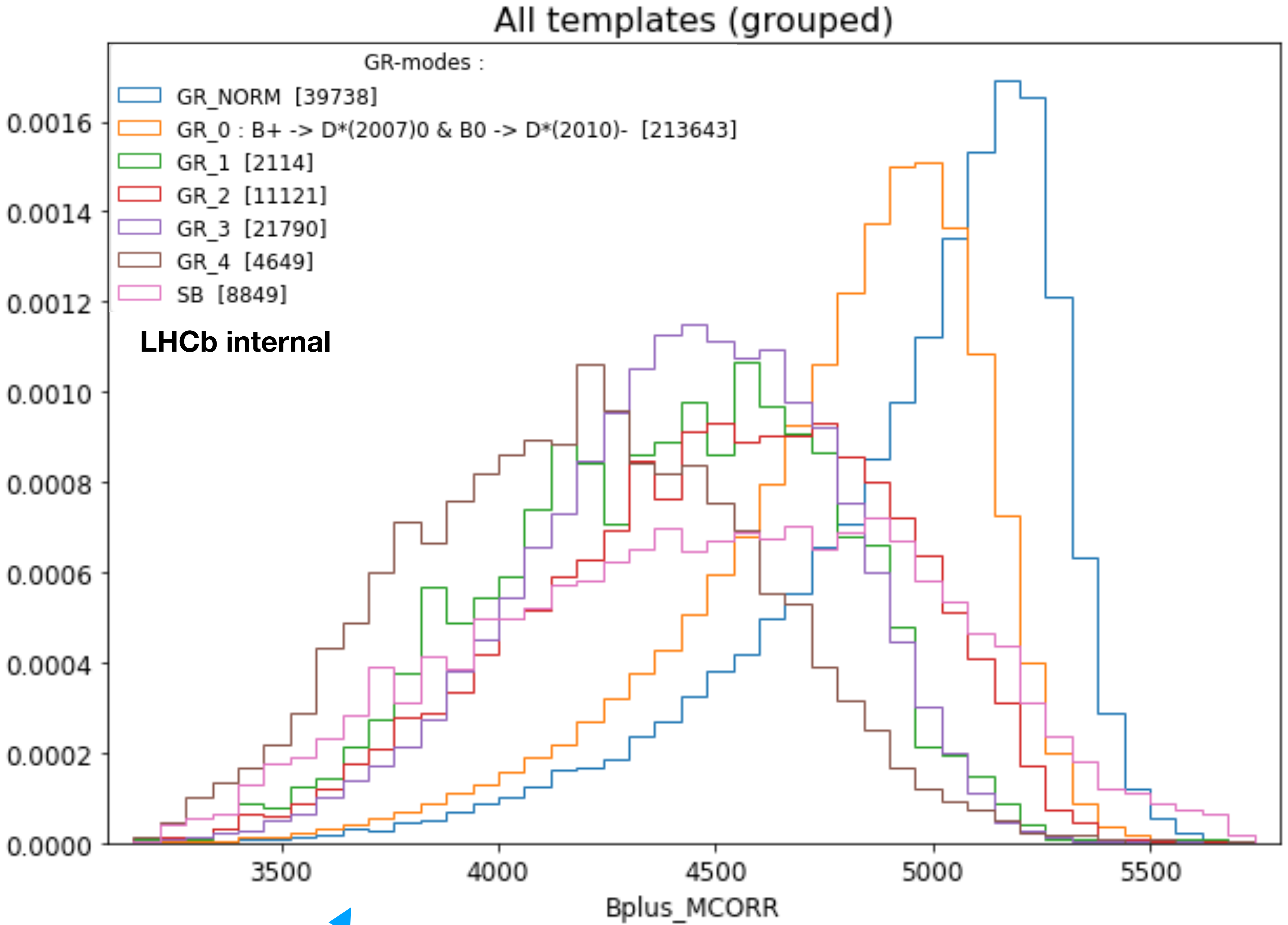
All modes in normalisation fit

- $B^+ \to \bar{D}^{(*,**)0}\mu^+\nu_\mu X$ **with** $\bar{D}^0 \to \pi^+\pi^-$ **cocktail (MC):**

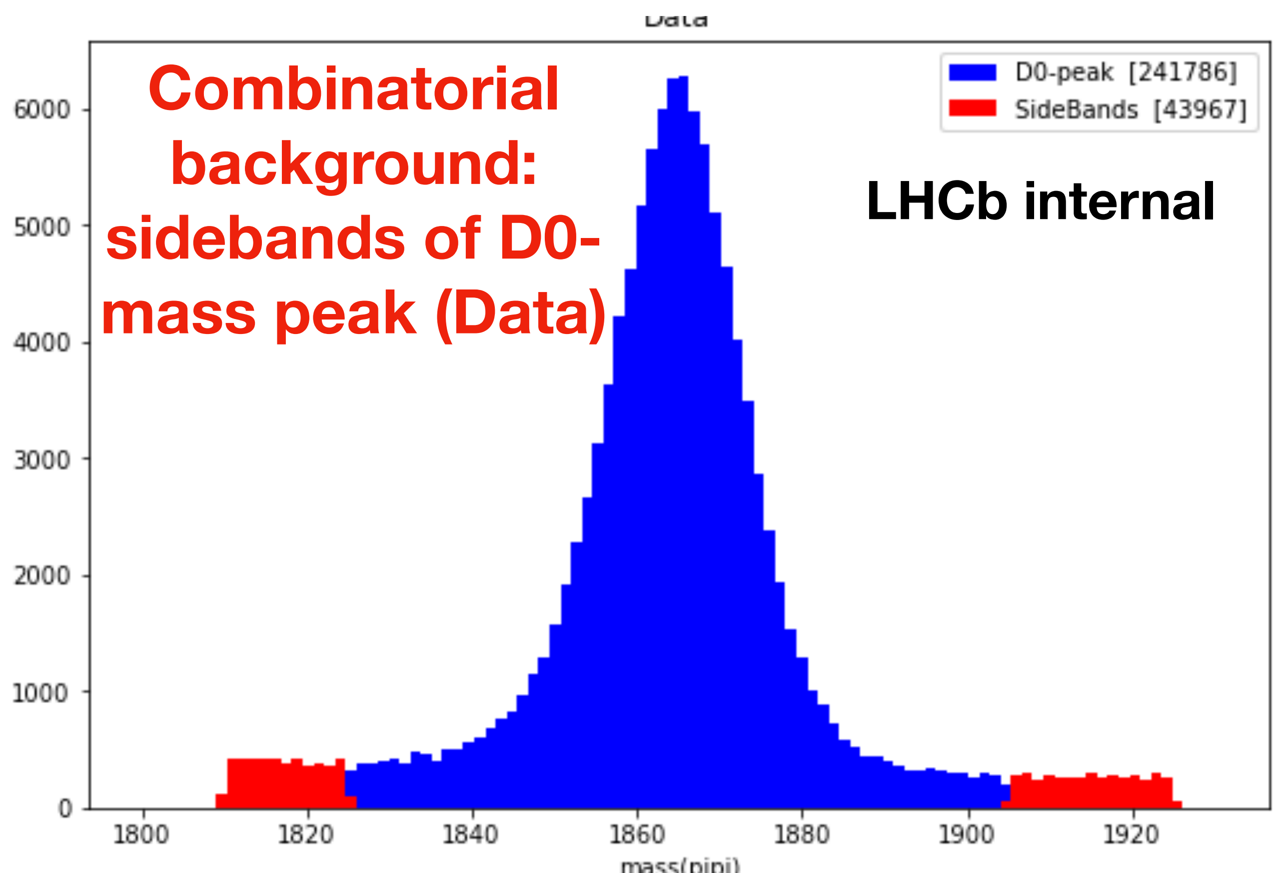
1. B+ -> D0 mu nu, D0->pi+pi-
2. B+ -> D0 mu nu pi0, D0->pi+pi-
3. B+ -> D*(2007)0 mu nu, D*(2007)0 -> D0 (pi0/gamma)
4. B+ -> D*(2007)0 mu nu pi0, D*(2007)0 -> D0 (pi0/gamma)
5. B+ -> D0*(2400)0 mu nu, D0*(2400)0 -> D0 pi0
6. B+ -> D1(2400)0 mu nu, D1(2400)0 -> D*(2010)+ pi-, D*(2010)+ -> D0 pi+
7. B+ -> D1(2400)0 mu nu, D1(2400)0 -> D*(2007)0 pi0, D*(2007)0 -> D0 (pi0/gamma)
8. B+ -> D1(2430)0 mu nu, D1(2430)0 -> D*(2010)+ pi-, D*(2010)+ -> D0 pi+
9. B+ -> D1(2430)0 mu nu, D1(2430)0 -> D*(2007)0 pi0, D*(2007)0 -> D0 (pi0/gamma)
10. B+ -> D2(2460)0 mu nu, D2(2460)0 -> D*(2010)+ pi-, D*(2010)+ -> D0 pi+
11. B+ -> D2(2460)0 mu nu, D2(2460)0 -> D0 pi0
12. B+ -> D2(2460)0 mu nu, D2(2460)0 -> D*(2007)0 pi0, D*(2007)0 -> D0 (pi0/gamma)
13. B+ -> D0 tau nu, tau -> mu nu_mu_bar nu_tau
14. B+ -> D*(2007)0 mu nu, tau -> mu nu_mu_bar nu_tau AND D*(2007)0 -> D0 (pi0/gamma)

- $B^0 \to D^{(*,**)-}\mu^+\nu_\mu X$ **with** $D^0 \to \pi^+\pi^-$ **cocktail (MC):**

1. B0 -> D*(2010)- mu+ nu, D*(2010)- -> D0_bar pi-
2. B0 -> D*(2010)- mu+ nu pi0, D*(2010)- -> D0_bar pi-
3. B0 -> D0*(2400)- mu+ nu, D*(2400)- -> D0_bar pi-
4. B0 -> D1(2420)- mu+ nu, D1(2420)- -> D*_bar(2007)0 pi-, D*_bar(2007)0 -> D0_bar (pi0/gamma)
5. B0 -> D1(2420)- mu+ nu, D1(2420)- -> D*(2010)- (pi0 / pi+ pi-), D*(2010)- -> D0_bar pi-
6. B0 -> D1(H)- mu+ nu, D1(H)- -> D_bar*(2007)0 pi-, D*_bar(2007)0 -> D0_bar (pi0/gamma)
7. B0 -> D1(H)- mu+ nu, D1(H)- -> D*(2010)- pi0, D*(2010)- -> D0_bar pi-
8. B0 -> D2*(2460)- mu+ nu, D2*(2460)- -> D0_bar pi+, D*_bar(2007)0 -> D0_bar (pi0/gamma)
9. B0 -> D2*(2460)- mu+ nu, D2*(2460)- -> D*_bar(2007)0 pi-, D*_bar(2007)0 -> D0_bar (pi0/gamma)
10. B0 -> D2*(2460)- mu+ nu, D2*(2460)- -> D*(2010)- pi0, D*(2010)- -> D0_bar pi-
11. B0 -> D0_bar pi- mu+ nu
12. B0 -> D*(2010)- tau nu, D*(2010)- -> D0_bar pi- AND tau -> mu nu_mu_bar nu_tau



Templates with similar m_{corr} shape are merged together (= GR-0, Gr-1,.. etc.)



All modes in signal fit

- **Signal and control channel:**

  (1) $B^+ \to \rho^0 \mu^+ \nu_\mu$

  (2) $B^0 \to J/\psi(\to \mu^+\mu^-)\rho^0$

- **Semileptonic V_{ub} backgrounds:**

  (3) $B^+ \to \omega^0 \mu^+ \nu_\mu$, $\omega \to \pi^+\pi^- X$

  (4) $B^+ \to \eta' \mu^+ \nu_\mu$, $\eta' \to \pi^+\pi^- X$

  - **Inclusive V_{ub} samples :**

    (5) $B^+ \to X_u \mu^+ \nu_\mu X$

    (6) $B^0 \to X_u \mu^+ \nu_\mu X$

- **Semileptonic V_{cb} decays where:**

  - $B^+ \to \bar{D}^{(*,**)0} \mu^+ \nu_\mu X$ with :

    (7) $\bar{D}^0 \to K_s^0 \pi^+ \pi^-$

    (8) $\bar{D}^0 \to K_s^0 \pi^+ \pi^- \pi^0$

    (9) $\bar{D}^0 \to \pi^+ \pi^- \pi^0 \pi^0$

    (10) $\bar{D}^0 \to \pi^+ \pi^- \pi^0$

    (11) $\bar{D}^0 \to K^+ \pi^- \pi^+ \pi^-$

  - $B^0 \to D^{(*,**)-} \mu^+ \nu_\mu X$ with :

    (12) $D^- \to \pi^+ \pi^- \pi^-$

Template histograms for the fit