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

Veronica S. Kirsebom & Michel De Cian

EPFL

SPS, Innsbruck
September, 2021
What is $B^+ \rightarrow \rho^0 \mu^+ \nu^\mu$?

- **Semileptonic decay:**

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

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

- **Measuring $B^+ \rightarrow \rho^0 \mu^+ \nu^\mu$ at LHCb:**

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

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

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 $\epsilon_R$ explanation.

---


Goals of this analysis

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

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

\[\Gamma(B^+ \to \rho^0[ \to \pi^+\pi^-]\mu^+\nu_\mu)\]

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 $q^2$, $\theta_l$, $\theta_V$ and $\phi$, and extract $\epsilon_R$.

Signal reconstruction

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

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

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

- Plots of $m_{\text{corr}}(B^+)$ and $m(\pi^+\pi^-)$ before any selection cuts:
What are the backgrounds?

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

- 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).

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

\[
\begin{align*}
\Gamma(b \to c) & \propto \frac{|V_{cb}|^2}{|V_{ub}|^2} \\
& \approx 100
\end{align*}
\]
Since our most prominent backgrounds often come with one or more additional tracks, we use an MVA trained on the isolatedness of the signal 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

Background rejection versus Signal efficiency

DNN ~3% better

MVA Method:
- PyKeras
- BDTAda
- MLPBNN

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

3.) **Combinatorial background from data**: 
   side-bands of the \( m(D^0) \)

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

\[
\begin{align*}
\text{Fit result:} & \quad S = 19586 \pm 434 \quad (~2.2\% \text{ relative error})
\end{align*}
\]
Template fit, we extract all templates from simulations of:

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

- **Semileptonic $V_{ub}$ backgrounds:**
  - $B^+ \rightarrow \omega^0 \mu^+ \nu_\mu$
  - $B^+ \rightarrow \eta' \mu^+ \nu_\mu$

Inclusive $V_{ub}$ samples

- **Semileptonic $V_{cb}$ decays where:**
  - $D \rightarrow \pi^+ \pi^- X$ with $X$ being 1-2 charged or neutral particles.

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

---

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

Room for improvements..
Conclusion and outlook

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

- Such a measurement can help us understand the $|V_{ub}|$ puzzle and a possible new physics explanation.
- We have developed a DNN that efficiently isolates signal against the dominant background of $b \rightarrow c$ 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 $|V_{ub}|$. 
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

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]:

\[
\mathcal{L}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{ub} \left( \bar{u}_\mu P_L b + \epsilon_R \bar{u}_\mu P_R b \right) (\bar{\nu}_\mu P L \ell) + \text{h.c.},
\]

| Decay | $|V_{ub}| \times 10^3$ | $\epsilon_R$ 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 | $1 - \epsilon_R$ |

<table>
<thead>
<tr>
<th>Decay</th>
<th>$B \times 10^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B \to \rho \ell \bar{\nu}$</td>
<td>1.97 ± 0.16 ($q^2 &lt; 12\text{ GeV}^2$)</td>
</tr>
<tr>
<td>$B \to \omega \ell \bar{\nu}$</td>
<td>0.61 ± 0.11 ($q^2 &lt; 12\text{ GeV}^2$)</td>
</tr>
</tbody>
</table>

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 $\{\cos \theta, \cos \theta_V\}$ phase space (integrated over $\phi$ and $q^2$).

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

\[
J_i = \frac{1}{N_i} \sum_{j=1}^{8} \sum_{k,l=1}^{4} \eta_{i,j} \eta_{i,k} \eta_{i,l} \left[ \lambda^{(j)} \otimes \theta^{(i)} \otimes \theta^{(k)} \right], \quad \langle P_{i,j} \rangle_{\text{bin}} = \frac{\int_{\Delta q^2} dq^2 J_i}{\int_{\Delta q^2} dq^2 J_j}.
\]

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**:

  \[ m_{\text{corr}}(B^+) = \sqrt{m_{\text{vis}}^2 + p_\perp^2 + p_\perp^2} \]

- 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.


**Derivation:** [https://lphe.epfl.ch/publications/theses/Lino_FerreiraLopes_MasterProject.pdf](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”
1. Diagonal cut

- We impose a requirement on the component of the $\rho^0$ momentum transverse to the $B^+$ flight direction, $P_{1B^+}(\rho^0)$, as a function of $m_{\text{corr}}(B^+)$ [4].

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

2. Diagonal cut

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

- 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 $B \to \rho \ell \nu$ (Fig. b) and $B \to \pi \ell \nu$ (Fig. a) due to their different spin structure.

\[ \frac{d\Gamma(B \to \rho \ell \nu)}{dq^2 d\cos\theta_W} = |V_{ub}|^2 \frac{G_F^2 P_0 q^2}{128\pi^3 M_B^2} \times \left[ \sin^2\theta_W |H_0|^2 + (1 - \cos \theta_W)^2 |H_2|^2 \right]. \]

- 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: $|P_B| = \frac{m_Y}{m_B} (P_Y)_z \sqrt{1 + \tan^2 \alpha}$, which is then used to boost the $\mu$ back in the $B$ rest frame.


Diagonal cuts

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

1.) Diagonal cut:

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

2.) Diagonal cut:

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

---

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**: $B^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \rho^0$ 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.

- The **main difference** between the two modes is the physics of the leptonic system: $W$-boson (weak force) vs. $\gamma$ (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**.

![Diagrams](image.png)

**The fit is good, and we obtain the s-weights**
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
DNN variable distributions

Input samples:
- Signal: $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$ 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.
- $\text{Mass}(B^+ + \text{track 1})$, M: invariant mass.
- $\text{mvaVal}(B^+ + \text{track 1})$, mva-ranking value.
- $\text{Mass}(\rho^0 + \text{track 1})$
- $\text{mvaVal}(\rho^0 + \text{track 1})$
- $\text{mvaVal}(\rho^0 + \text{track 1} & 2)$
$\pi^+\pi^-$-mass distribution after DNN cut
All modes in normalisation fit

- **$B^+ \rightarrow D^{(*)\pm\pm} / \mu^+ \nu \mu X$ with $D^0 \rightarrow \pi^+\pi^-$ cocktail (MC):**
  1. $B^+ \rightarrow D^0 \mu^+ \nu, D^0 \rightarrow \pi^+\pi^-$
  2. $B^+ \rightarrow D^0 \mu^+ \nu \pi^0, D^0 \rightarrow \pi^+\pi^-$
  3. $B^+ \rightarrow D^*\!(2007)^0 \mu^+ \nu, D^*\!(2007)^0 \rightarrow D^0 (\pi^0/\gamma)$
  4. $B^+ \rightarrow D^*\!(2007)^0 \mu^+ \nu \pi^0, D^*\!(2007)^0 \rightarrow D^0 (\pi^0/\gamma)$
  5. $B^+ \rightarrow D^0(2400)^0 \mu^+ \nu, D^0(2400)^0 \rightarrow D^0 \pi^0$
  6. $B^+ \rightarrow D^*\!(2007)^0 \mu^+ \nu, D^*\!(2007)^0 \rightarrow D^0 \pi^+\pi^-$
  7. $B^+ \rightarrow D(2430)^0 \mu^+ \nu, D(2430)^0 \rightarrow D^*\!(2007)^0 \pi^0, D^*\!(2007)^0 \rightarrow D^0 (\pi^0/\gamma)$
  8. $B^+ \rightarrow D(2430)^0 \mu^+ \nu \pi^0, D(2430)^0 \rightarrow D^*\!(2007)^0 \pi^0, D^*\!(2007)^0 \rightarrow D^0 (\pi^0/\gamma)$
  9. $B^+ \rightarrow D(2460)^0 \mu^+ \nu, D(2460)^0 \rightarrow D^*\!(2010)^0 \pi^0, D^*\!(2010)^0 \rightarrow D^0 (\pi^0/\gamma)$
  10. $B^+ \rightarrow D(2460)^0 \mu^+ \nu \pi^0, D(2460)^0 \rightarrow D^*\!(2010)^0 \pi^0, D^*\!(2010)^0 \rightarrow D^0 (\pi^0/\gamma)$
  11. $B^+ \rightarrow D(2460)^0 \mu^+ \nu, D(2460)^0 \rightarrow D^0 \pi^0$
  12. $B^+ \rightarrow D(2460)^0 \mu^+ \nu \pi^0, D(2460)^0 \rightarrow D^*\!(2007)^0 \pi^0, D^*\!(2007)^0 \rightarrow D^0 (\pi^0/\gamma)$
  13. $B^+ \rightarrow D(2460)^0 \mu^+ \nu, D(2460)^0 \rightarrow D^0 (\pi^0/\gamma)$
  14. $B^+ \rightarrow D(2460)^0 \mu^+ \nu \pi^0, D(2460)^0 \rightarrow D^*\!(2007)^0 \pi^0, D^*\!(2007)^0 \rightarrow D^0 (\pi^0/\gamma)$
  15. $B^+ \rightarrow D(2460)^0 \mu^+ \nu, D(2460)^0 \rightarrow D^0 (\pi^0/\gamma)$

- **$B^0 \rightarrow D^{(*)\pm\pm} / \mu^+ \nu \mu X$ with $D^0 \rightarrow \pi^+\pi^-$ cocktail (MC):**
  1. $B^0 \rightarrow D^\mp\!(2010)^0 \mu^+ \nu, D^\mp\!(2010)^0 \rightarrow D^0 \pi^+\pi^-$
  2. $B^0 \rightarrow D^\mp\!(2010)^0 \mu^+ \nu \pi^0, D^\mp\!(2010)^0 \rightarrow D^0 \pi^+\pi^-$
  3. $B^0 \rightarrow D^\mp\!(2400)^0 \mu^+ \nu, D^\mp\!(2400)^0 \rightarrow D^0 \pi^+\pi^-$
  4. $B^0 \rightarrow D^\mp\!(2420)^0 \mu^+ \nu, D^\mp\!(2420)^0 \rightarrow D^\mp\!\!(2007)^0 \pi^0 / \pi^+\pi^-, D^\mp\!\!(2007)^0 \rightarrow D^0 \pi^+\pi^-$
  5. $B^0 \rightarrow D^\mp\!\!(2007)^0 \mu^+ \nu, D^\mp\!\!(2007)^0 \rightarrow D^0 \pi^+\pi^-$
  6. $B^0 \rightarrow D^\mp\!\!(2007)^0 \mu^+ \nu \pi^0, D^\mp\!\!(2007)^0 \rightarrow D^0 \pi^+\pi^-$
  7. $B^0 \rightarrow D^\mp\!\!(2010)^+, D^\mp\!\!(2010)^+ \rightarrow D^0 \pi^+\pi^-$
  8. $B^0 \rightarrow D^\mp\!\!(2460)^0 \mu^+ \nu, D^\mp\!\!(2460)^0 \rightarrow D^0 \pi^+\pi^-$
  9. $B^0 \rightarrow D^\mp\!\!(2460)^0 \mu^+ \nu \pi^0, D^\mp\!\!(2460)^0 \rightarrow D^0 \pi^+\pi^-$

Combinatorial background: sidebands of $D^0$-mass peak (Data)

Templates with similar $m_{corr}$ shape are merged together (= GR-0, Gr-1, etc.)
All modes in signal fit

• Signal and control channel:
  - (1) $B^+ \rightarrow \rho^0 \mu^+ \nu_\mu$
  - (2) $B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \rho^0$

• Semileptonic $V_{ub}$ backgrounds:
  - (3) $B^+ \rightarrow \omega^0 \mu^+ \nu_\mu$, $\omega \rightarrow \pi^+ \pi^- X$
  - (4) $B^+ \rightarrow \eta \mu^+ \nu_\mu$, $\eta' \rightarrow \pi^+ \pi^- X$
    • Inclusive $V_{ub}$ samples:
      - (5) $B^+ \rightarrow X_{ub} \mu^+ \nu_\mu X$
      - (6) $B^0 \rightarrow X_{ub} \mu^+ \nu_\mu X$

• Semileptonic $V_{cb}$ decays where:
  - $B^+ \rightarrow \bar{D}^{(*) \rightarrow \pi \pi \mu}$ with:
    - (7) $\bar{D}^0 \rightarrow K^0_s \pi^+ \pi^-$
    - (8) $\bar{D}^0 \rightarrow K^0_s \pi^+ \pi^- \pi^0$
    - (9) $\bar{D}^0 \rightarrow \pi^+ \pi^- \pi^0 \pi^0$
    - (10) $\bar{D}^0 \rightarrow \pi^+ \pi^- \pi^0$
    - (11) $\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$
  - $B^0 \rightarrow D^{(*) \rightarrow \pi \pi \mu}$ with:
    - (12) $D^- \rightarrow \pi^+ \pi^- \pi^-$