HNL production from B-meson and W-boson decays at LHCb

Rate estimates at $\sqrt{s} = 13$ TeV and comparison of production channels

HNL sensitivity working note

October 23, 2026

Abstract

We collect the ingredients needed to estimate heavy-neutral-lepton (HNL, N) production at LHCb from B-hadron decays, and compare them with the $W \to \mu N$ production mode used in the existing sensitivity studies. Section 1 summarises the 13 TeV production cross sections of the four B mesons (and the Λ_b) in the LHCb acceptance and contrasts them with W production. Section 2 derives the per-channel N-production branching ratios as functions of m_N (leptonic two-body decays with helicity enhancement, semileptonic decays with a q^2 retention factor, and $W \to \mu N$). Section 3 assembles these into produced-N counts for an integrated luminosity of 30 fb^{-1} at $|U_{\mu N}|^2 = 1$, for $m_N = 1$–8 GeV. Section 4 analyses, for each major production $\times$ decay signature, the dominant backgrounds and the selections that suppress them. Section 5 discusses the hierarchy of channels, the mass ranges they cover, and the lifetime caveat that matters for a displaced (muon-detector) search. All N rates scale as $\propto |U_{\mu N}|^2$ unless stated otherwise.

1 Production of B mesons and W bosons at the LHC

1.1 The four B mesons

There are exactly four ground-state open-beauty mesons, distinguished by the light (or charm) antiquark bound to the b quark:

meson	quark content	mass [GeV]	lifetime [ps]	remark
B^+	$b\bar{u}$	5.279	1.64	only light B with 2-body leptonic decay
B^0	$b\bar{d}$	5.280	1.52	neutral: semileptonic only
B_s^0	$b\bar{s}$	5.367	1.51	neutral: semileptonic only
B_c^+	$b\bar{c}$	6.275	0.51	heaviest; largest HNL mass reach

Each has an antiparticle (B^-, $\bar{B}^0$, $\bar{B}_s^0$, B_c^-). Everything else in the b-hadron spectrum is either a baryon (Λ_b, Ξ_b, Σ_b, Ω_b), bottomonium ($b\bar{b}$: Υ family, no open flavour), or an excited state (B^{**}, $B_c(2S)$, …) that decays promptly to one of the four ground states above. All open-beauty production therefore funnels into these four species (plus the baryons) before any weak decay of interest takes place.

1.2 Cross sections at $\sqrt{s} = 13$ TeV

The inclusive $b\bar{b}$ production cross section is

$$\sigma(pp \to b\bar{b}\,X) \simeq 560\ \mu\text{b} \quad (4\pi), \tag{1}$$

of which the LHCb measurement (semileptonic b-hadron decays, [1])

$$\sigma(pp \to b\bar{b}\,X,\ 2 < \eta < 5) = 154.3 \pm 1.5 \pm 14.3\ \mu\text{b} \tag{2}$$

lands in the forward acceptance — about 28% of the total, reflecting the soft, forward-peaked nature of QCD heavy-flavour production. The split into species is governed by the fragmentation fractions measured by LHCb at 13 TeV ($p_T \in [4, 25]$ GeV, $2 < \eta < 5$) [2, 3]:

$$\frac{f_s}{f_u + f_d} = 0.122 \pm 0.006, \qquad \frac{f_{\Lambda_b}}{f_u + f_d} = 0.259 \pm 0.018, \qquad \frac{f_c}{f_u + f_d} = (3.78 \pm 0.89) \times 10^{-3}, \quad (3)$$

which, with isospin symmetry $f_u = f_d$ and normalising to unity, give $f_u = f_d \simeq 0.36$, $f_s \simeq 0.088$, $f_{\Lambda_b} \simeq 0.19$, $f_c \simeq 1.4 \times 10^{-3}$. The resulting per-species cross sections and parent counts at 30 fb^{-1} are:

species	fraction	σ in $2 < \eta < 5$ [μb]	N at 30 fb^{-1}
B^+	0.36	55	1.65×10^{12}
B^0	0.36	55	1.65×10^{12}
B_s^0	0.088	14	4.2×10^{11}
Λ_b (+ other baryons)	0.19	29	8.7×10^{11}
B_c^+	1.4×10^{-3}	0.2	6.0×10^{9}

Table 1: B-hadron production in the LHCb acceptance at 13 TeV. Caveats: f_{Λ_b} is strongly p_T-dependent; f_c carries a ∼25% theory uncertainty from the BR($B_c \to J/\psi\,\mu\nu$) input used to extract it.

1.3 W production

Inclusive W production at 13 TeV is, at NNLO [4],

$$\sigma(W^+) \simeq 11.7 \text{ nb}, \qquad \sigma(W^-) \simeq 8.6 \text{ nb}, \qquad \sigma(W^+ + W^-) \simeq 20.3 \text{ nb}, \quad (4)$$

in agreement with ATLAS/CMS measurements at the ∼2% level. Multiplying by the leptonic branching ratio BR($W \to \mu\nu$) = 10.86%,

$$\sigma(W \to \mu\nu) \simeq 2.2 \text{ nb} \quad \Longrightarrow \quad N(W \to \mu\nu)\big|_{30 \text{ fb}^{-1}} \simeq 6.6 \times 10^7 \ (4\pi), \quad (5)$$

of which $\mathcal{O}$(15–20%) falls in the LHCb η acceptance. (This is the normalisation used in the existing analysis chain, where $K_{\text{NNLO}} = 20.08/9.691$ rescales the LO MadGraph cross section to $\sigma_{\text{NNLO}} = 20.08$ nb.)

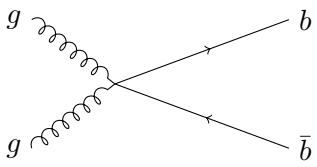
1.4 Why the $b\bar{b}$ rate dwarfs the W rate

The ratio

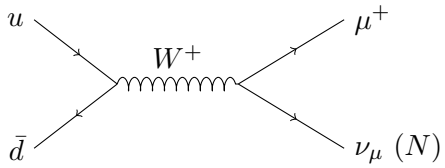
$$\frac{\sigma(b\bar{b})}{\sigma(W)} \simeq \frac{560\ \mu\text{b}}{20.3\ \text{nb}} \simeq 2.8 \times 10^4 \quad (6)$$

has a simple dynamical origin:

- **Coupling and order.** $b\bar{b}$ production is a strong process, $\mathcal{O}(\alpha_s^2)$ at Born level ($gg \to b\bar{b}$, $q\bar{q} \to b\bar{b}$, plus the dominant-at-LHC t-channel topologies: flavour excitation $qg \to qb\bar{b}$ and gluon splitting $g \to b\bar{b}$). W production is electroweak Drell–Yan, $q\bar{q}' \to W$, of order α^2.
- **Mass scale.** The hard scale for $b\bar{b}$ is set by $2m_b \simeq 9.4$ GeV $\ll \sqrt{s}$, in a region where gluon PDFs are enormous; the W sits at $m_W = 80$ GeV where the quark–antiquark luminosity is far smaller. The parton-luminosity penalty is worth roughly an order of magnitude on its own.



(a) $gg \to b\bar{b}$, $\mathcal{O}(\alpha_s^2)$



(b) $u\bar{d} \to W^+ \to \mu^+\nu_\mu$, $\mathcal{O}(\alpha^2)$

Figure 1: Born topologies behind the two production scales: (a) strong $b\bar{b}$ pair production (the dominant LHC topologies also include flavour excitation and gluon splitting); (b) electroweak Drell–Yan W production.

- **Phase space/topology.** $b\bar{b}$ is pair production with a large soft and collinear enhancement — hence its strong forward peaking that makes LHCb's $2 < \eta < 5$ arm capture ∼28% of it. Drell–Yan W's are single production, comparatively central, with a forward component driven by valence quarks at large x hitting the low-x sea.

The same logic fixes the hierarchy among the B species: $f_u = f_d$ by isospin; $f_s/f_d \simeq 0.25$ reflects the strangeness penalty in hadronisation; the baryon fraction is substantial at forward p_T; and B_c is suppressed by $\sim 260\times$ relative to B^+ because it requires *two* heavy-quark pairs (b and $\bar{c}$) produced close enough in phase space to bind.

2 HNL production mechanisms and per-channel rates

A muon-coupled HNL inherits the charged-current coupling of ν_μ through the mixing $U_{\mu N}$. Every b-hadron (or W) decay with a ν_μ in the final state therefore has an N analogue with

$$\mathrm{BR}(\mathrm{parent} \to X\,\mu N) = |U_{\mu N}|^2 \times \mathrm{BR}(\mathrm{parent} \to X\,\mu\nu) \times R(m_N), \tag{7}$$

where $R(m_N)$ collects all mass effects: phase space, helicity factors, and — in the two-body leptonic case — a large *enhancement*. We now treat each class.

2.1 Leptonic two-body decays $B^+ \to \mu^+ N$

For a spin-0 meson P of mass m_P decaying to $\ell^+ N$ (writing $r_i = m_i^2/m_P^2$),

$$\Gamma(P \to \ell N) = \frac{G_F^2 f_P^2 |V_{q_1 q_2}|^2 |U_{\mu N}|^2}{8\pi}\, m_P^3\, \lambda^{1/2}(1, r_\ell, r_N)\left[r_\ell + r_N - (r_\ell - r_N)^2\right], \tag{8}$$

with $\lambda(a,b,c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc$ the Källén function. The key physics is in the last bracket. In the SM mode $P \to \ell\nu$ only the *lepton* can carry the required wrong helicity, so the rate is proportional to $r_\ell(1-r_\ell)^2$: the famous helicity suppression ($\mathrm{BR}(B^+ \to \mu^+\nu) = 5.3 \times 10^{-7}$, $\mathrm{BR}(B^+ \to e^+\nu) \sim 10^{-11}$). For a massive N there is a second wrong-helicity option, proportional to m_N, and the suppression is *lifted*. Relative to the massless-ν mode (for $r_\ell \to 0$):

$$R(m_N) \;=\; \frac{\lambda^{1/2}(1, r_\ell, r_N)\left[r_\ell + r_N - (r_\ell - r_N)^2\right]}{r_\ell(1-r_\ell)^2} \;\simeq\; \frac{m_N^2}{m_\mu^2}\left(1 - \frac{m_N^2}{m_P^2}\right)^2, \tag{9}$$

i.e. an enhancement by $m_N^2/m_\mu^2 \simeq$ 90–3600 over the relevant mass range, cut off only by the two-body phase space at $m_N \to m_P$. The resulting branching ratios per unit $|U_{\mu N}|^2$ are listed in Table 2; $\mathrm{BR}(B^+ \to \mu^+ N)$ peaks at $\simeq 2 \times 10^{-4}\,|U_{\mu N}|^2$ around $m_N \simeq 3$ GeV — two to three orders of magnitude above the naive $\mathrm{BR}(B \to \mu\nu) \times |U_{\mu N}|^2$ guess. Note also:

- $\mathrm{BR}(B^+ \to e^+ N) \simeq \mathrm{BR}(B^+ \to \mu^+ N)$ at $m_N \gg m_\mu$: the m_N^2/m_e^2 enhancement exactly compensates the tinier SM rate. The channel probes the same $|U|^2$ whichever lepton flavour is mixed.

- $B^+ \to \tau^+ N$ probes $|U_{\tau N}|^2$ for $m_N < m_B - m_\tau = 3.5$ GeV.
- $B_c^+ \to \mu^+ N$ enjoys both a larger phase space and a larger SM normalisation, $\mathrm{BR}(B_c \to \mu\nu) \simeq 4 \times 10^{-4}$ (larger f_{B_c}): $\mathrm{BR}(B_c \to \mu N)$ reaches $\sim 0.2\,|U_{\mu N}|^2$ at $m_N \simeq 4$ GeV and stays open to $m_N \simeq m_{B_c} = 6.27$ GeV — the largest reach of any B decay. Only B^+ and B_c^+ (charged) have two-body leptonic decays.

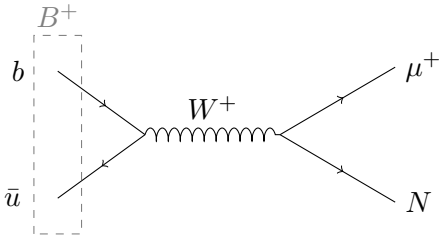


Figure 2: Leptonic production $B^+ \to \mu^+ N$: the $b\bar{u}$ pair annihilates into an off-shell W; the massive N supplies the wrong helicity, lifting the m_μ^2 suppression of the SM mode $B^+ \to \mu^+\nu$.

2.2 Semileptonic decays $B \to X\,\mu N$

Semileptonic decays carry no helicity argument either way: the N simply replaces the ν with a phase-space penalty. Writing q^2 for the invariant mass squared of the μN system,

$$\mathrm{BR}(B \to X\mu N;\, m_N) = |U_{\mu N}|^2 \times \mathrm{BR}(B \to X\mu\nu) \times R(m_N), \tag{10}$$

$$R(m_N) = \frac{\displaystyle\int_{m_N^2}^{q^2_{\max}} dq^2\, \frac{d\Gamma_0}{dq^2} \left(1 - \frac{m_N^2}{q^2}\right)^2 \left(1 + \frac{m_N^2}{2q^2}\right)}{\displaystyle\int_0^{q^2_{\max}} dq^2\, \frac{d\Gamma_0}{dq^2}}, \tag{11}$$

with $q^2_{\max} = (m_B - m_X)^2$ and $\rho(m_N, q^2) = (1 - m_N^2/q^2)^2(1 + m_N^2/2q^2)$ the standard massive-lepton factor. For the estimates below we model the massless-ν spectrum as

$$\frac{d\Gamma_0}{dq^2} \propto \lambda^{3/2}\big(m_B^2, m_X^2, q^2\big)\,, \tag{12}$$

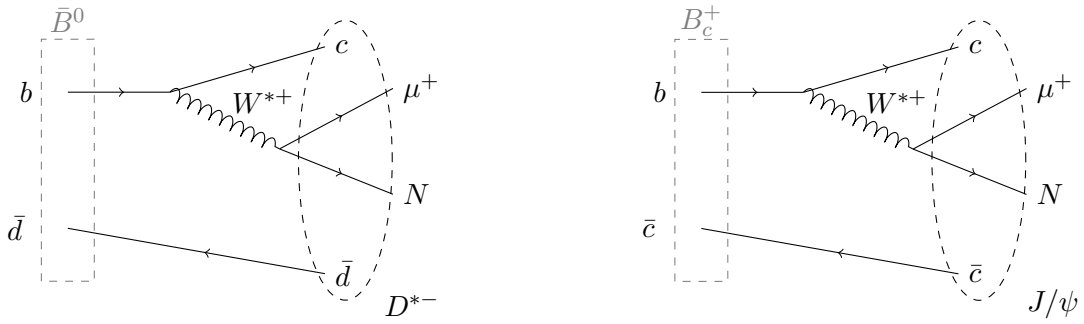
the $B \to P\ell\nu$ shape with a constant form factor; this reproduces the measured $B \to D^{(*)}\ell\nu$ spectra to ∼30–50% in the integrated retention factors, which is adequate for rate scouting. The resulting $R(m_N)$ (Table 2) have two important features:

- The spectrum is soft (it falls steeply with q^2), so the retention drops quickly: $R \simeq 0.38$–0.40 at $m_N = 1$ GeV but only $\simeq 0.04$–0.05 at $m_N = 2$ GeV for the $D^{(*)}$ modes.
- Each channel closes sharply at $m_N = m_B - m_X$: $B \to D^*\mu N$ at 3.27 GeV, $B \to D\mu N$ at 3.41, $B_s \to D_s\mu N$ at 3.40, $\Lambda_b \to \Lambda_c\mu N$ at 3.33, $B_c \to J/\psi\,\mu N$ at 3.18. The $b \to u$ modes $B \to \pi\mu N$ and $B_s \to K\mu N$ survive to 5.14 and 4.87 GeV respectively but with tiny normalisations ($\mathrm{BR}(B \to \pi\ell\nu) \simeq 1.5 \times 10^{-4}$).

2.3 $W \to \mu N$ (reference channel)

The production mode of the existing sensitivity chain is the on-shell W (Fig. 1b). For a two-body decay of the W,

$$\mathrm{BR}(W \to \mu N) = |U_{\mu N}|^2 \times \mathrm{BR}(W \to \mu\nu) \times (1 - r_N)^2 \left(1 + \frac{r_N}{2}\right), \qquad r_N = \frac{m_N^2}{m_W^2}, \tag{13}$$



(a) $\bar{B}^0 \to D^{*-}\mu^+ N$ (b) $B_c^+ \to J/\psi\,\mu^+ N$

Figure 3: Semileptonic production: the $b \to c$ charged current produces the $\mu^+ N$ pair while the spectator quark dresses the recoiling charm into a $D^{(*)}$ (a), D_s, or J/ψ (b). No helicity enhancement; the rate is $\mathrm{BR}(B \to X\mu\nu) \times |U_{\mu N}|^2 \times R(m_N)$ of Eq. (11).

with $\mathrm{BR}(W \to \mu\nu) = 10.86\%$. Since $m_W = 80.4$ GeV, the phase-space factor is ≥ 0.98 across the whole $m_N = 1$–8 GeV range considered here: the W channel is essentially mass-independent in this window, with $\mathrm{BR}(W \to \mu N) \simeq 0.108\,|U_{\mu N}|^2$.

2.4 The decay side (why these are "decay chains")

The same mixing opens the N decay, dominated at these masses by the charged-current modes $N \to \mu^\pm q\bar{q}'$ (hadronising e.g. to $\mu^\pm\pi^\mp$, $\mu^\pm\rho^\mp$) and the neutral-current/leptonic modes (Fig. 4). The total width scales as

$$\Gamma(N) \simeq |U_{\mu N}|^2 \times \Gamma_0(m_N), \qquad \Gamma_0(m_N) \sim \frac{G_F^2 m_N^5}{192\pi^3} \times \mathcal{O}(5\text{–}10), \tag{14}$$

so at $|U_{\mu N}|^2 = 1$ the lifetime is negligible ($c\tau \sim 10^{-7}$ mm at $m_N = 2$ GeV) and every produced N decays promptly; the long lifetimes relevant for displaced searches appear at $|U_{\mu N}|^2 \lesssim 10^{-8}$–$10^{-10}$, where the production rates of Sec. 3 must simply be rescaled by $|U_{\mu N}|^2$. For the production-rate comparison of this note, $|U_{\mu N}|^2 = 1$ is the natural normalisation: it factors out of every channel identically.

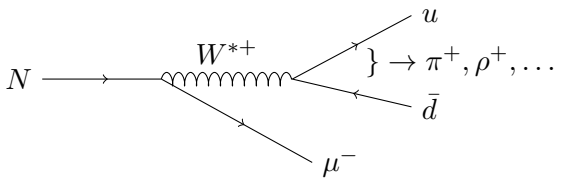


Figure 4: Dominant N decay topology at GeV-scale masses (Majorana N shown; for Dirac N the charge-conjugate final state applies). The displaced $\mu^\pm\pi^\mp$ signature used in existing LHCb/Belle searches comes from this chain.

3 Produced HNL counts at 30 fb^{-1}, $|U_{\mu N}|^2 = 1$

Table 2 collects the per-channel branching ratios of Secs. 2.1–2.3 as functions of m_N, and Table 3 multiplies them by the parent production of Sec. 1 to give the number of N's produced per unit $|U_{\mu N}|^2$ for an integrated luminosity of 30 fb^{-1}. Inputs: masses and lifetimes from the PDG [5]; $\mathrm{BR}(B^+ \to \mu^+\nu) = 5.3 \times 10^{-7}$, $\mathrm{BR}(B_c^+ \to \mu^+\nu) \simeq 4 \times 10^{-4}$ (theory), $\mathrm{BR}(\bar{B}^0 \to D^{*-}\mu^+\nu) = 5.05\%$, $\mathrm{BR}(B^+ \to \bar{D}^{*0}\mu^+\nu) = 5.58\%$ (merged below with the D modes to a common $\mathrm{BR}_0 = 5.3\%$), $\mathrm{BR}(B \to D\mu\nu) = 2.3\%$, $\mathrm{BR}(B \to \pi\mu\nu) = 1.5\times10^{-4}$ ($+B^+ \to \pi^0\mu\nu = 0.78\times10^{-4}$),

$\mathrm{BR}(B_s \to D_s\mu\nu) = 2.3\%$, $\mathrm{BR}(B_s \to K^-\mu\nu) = 1.9 \times 10^{-4}$, $\mathrm{BR}(B_c \to J/\psi\,\mu\nu) \simeq 1.5\%$ (theory), $\mathrm{BR}(\Lambda_b \to \Lambda_c\mu\nu) \simeq 6\%$, $\mathrm{BR}(W \to \mu\nu) = 10.86\%$.

channel	BR(channel)/$\lvert U_{\mu N}\rvert^2$ at m_N [GeV]							
	1	2	3	4	5	6	7	8
$W \to \mu N$ (Fig. 1b)	0.1086	0.1085	0.1084	0.1082	0.1080	0.1077	0.1074	0.1070
$B^+ \to \mu^+ N$ (Fig. 2)	4.5×10^{-5}	1.4×10^{-4}	2.0×10^{-4}	1.4×10^{-4}	1.2×10^{-5}	— closed —		
$B \to D^*\mu N$	2.0×10^{-2}	2.0×10^{-3}	2.3×10^{-6}	— closed ($m_N > 3.27$) —				
$B \to D\mu N$	9.2×10^{-3}	1.1×10^{-3}	5.1×10^{-6}	— closed ($m_N > 3.41$) —				
$B \to \pi\mu N$	1.2×10^{-4}	3.6×10^{-5}	5.4×10^{-6}	2.0×10^{-7}	5.5×10^{-12}	— closed —		
$B_s \to D_s\mu N$	9.2×10^{-3}	1.1×10^{-3}	4.6×10^{-6}	— closed ($m_N > 3.40$) —				
$B_s \to K^-\mu N$	1.0×10^{-4}	3.0×10^{-5}	4.3×10^{-6}	1.1×10^{-7}	— closed ($m_N > 4.87$) —			
$B_c^+ \to \mu^+ N$	3.4×10^{-2}	1.2×10^{-1}	1.9×10^{-1}	2.0×10^{-1}	1.2×10^{-1}	8.8×10^{-3}	— closed —	
$B_c \to J/\psi\,\mu N$	5.5×10^{-3}	4.9×10^{-4}	1.3×10^{-7}	— closed ($m_N > 3.18$) —				
$\Lambda_b \to \Lambda_c\mu N$	2.3×10^{-2}	2.6×10^{-3}	6.2×10^{-6}	— closed ($m_N > 3.33$) —				

Table 2: HNL-production branching ratios per unit $|U_{\mu N}|^2$ for the channels of Sec. 2. Leptonic rows use Eq. (8); semileptonic rows use Eq. (11) with the $\lambda^{3/2}$ spectrum model (∼30–50% accuracy in the retention factor); the W row uses Eq. (13). "Closed" marks the two-body phase-space boundary $m_N \geq m_{\text{parent}} - m_X$; the approach to it is in practice faster than the boundary, cf. the $B \to D^*$ row at $m_N = 3$ GeV.

channel	parents	N produced at 30 fb^{-1}, $\lvert U_{\mu N}\rvert^2 = 1$							
	at 30 fb^{-1}	1	2	3	4	5	6	7	8
$W \to \mu N$	6.6×10^7 (4π)	$\simeq 6.6 \times 10^7$ at every mass in the table							
$B^+ \to \mu^+ N$	1.65×10^{12}	7.4×10^7	2.3×10^8	3.2×10^8	2.3×10^8	2.0×10^7	—	—	—
$B \to D^*\mu N$	3.3×10^{12}	6.6×10^{10}	6.6×10^9	7.5×10^6	—	—	—	—	—
$B \to D\mu N$	3.3×10^{12}	3.0×10^{10}	3.7×10^9	1.7×10^7	—	—	—	—	—
$B \to \pi\mu N$	3.3×10^{12}	4.0×10^8	1.2×10^8	1.8×10^7	6.7×10^5	18	—	—	—
$B_s \to D_s\mu N$	4.2×10^{11}	3.8×10^9	4.6×10^8	1.9×10^6	—	—	—	—	—
$B_s \to K^-\mu N$	4.2×10^{11}	4.4×10^7	1.3×10^7	1.8×10^6	4.8×10^4	—	—	—	—
$B_c^+ \to \mu^+ N$	6.0×10^9	2.1×10^8	7.0×10^8	1.2×10^9	1.2×10^9	7.2×10^8	5.3×10^7	—	—
$B_c \to J/\psi\,\mu N$	6.0×10^9	3.3×10^7	3.0×10^6	7.8×10^2	—	—	—	—	—
$\Lambda_b \to \Lambda_c\mu N$	8.7×10^{11}	2.0×10^{10}	2.3×10^9	5.4×10^6	—	—	—	—	—
sum, all B channels		1.2×10^{11}	1.4×10^{10}	1.5×10^9	1.4×10^9	7.4×10^8	5.3×10^7	—	—

Table 3: Number of HNLs produced per unit $|U_{\mu N}|^2$ at 30 fb^{-1}, obtained by multiplying Table 2 by the parent populations of Table 1 (in $2 < \eta < 5$; the W row is 4π production — the η-filtered fraction is ∼15–20%). $B \to D^{(*)}$ and $B \to \pi$ rows merge B^+ and B^0 (identical kinematics within MeV). Every entry scales linearly with $|U_{\mu N}|^2$; note however that the N lifetime scales as $1/|U_{\mu N}|^2$, so the interesting long-lived regime sits at $|U_{\mu N}|^2 \ll 1$ (Sec. 5).

Reading the tables (and Fig. 5). At $|U_{\mu N}|^2 = 1$ the B-decay production exceeds W-decay production by a factor ~ 2000 at $m_N = 1$ GeV, ~ 200 at 2 GeV, ~ 20 at 3–5 GeV, and falls to parity around $m_N \simeq 6$ GeV; above the B_c threshold ($m_N > 6.2$ GeV) the W is the only source. The channel hierarchy inside the B sector changes qualitatively with mass: semileptonic $B \to D^{(*)}\mu N$ dominates up to $m_N \simeq 2$ GeV, the leptonic $B^+ \to \mu^+ N$ (helicity-enhanced) dominates from $m_N \simeq 2.5$ GeV to the B mass, and $B_c^+ \to \mu^+ N$ alone covers 5.2–6.2 GeV. B_c is remarkably competitive with B^+ in the leptonic row despite its 260× smaller production rate, because its two-body BR is $\sim 10^3$ times larger.

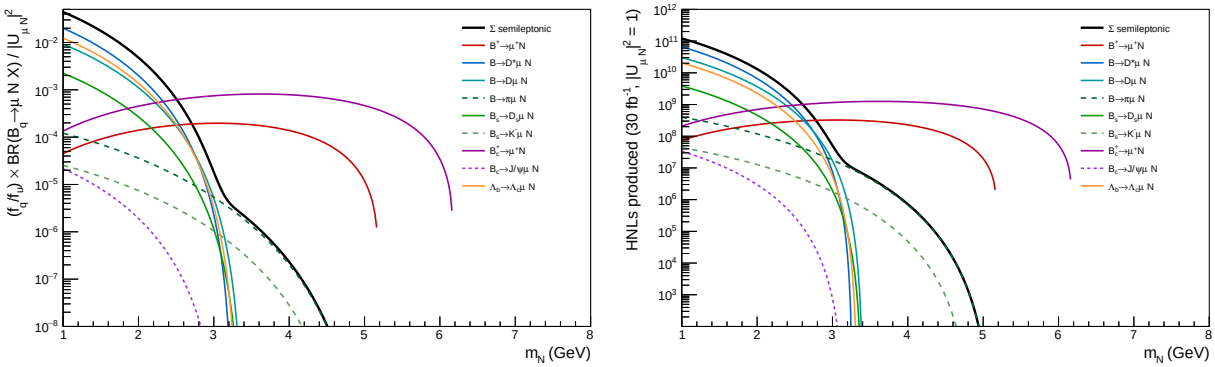


Figure 5: Left: HNL-production branching ratio per unit $|U_{\mu N}|^2$, weighted by the fragmentation fraction f_q/f_u — i.e. Table 2 multiplied by f_q/f_u — as a continuous function of m_N. Right: produced HNL counts at 30 fb^{-1} for $|U_{\mu N}|^2 = 1$ (Table 3). Curves end where the channel closes kinematically. Produced by `scripts/plot_rates.py` (python/ROOT), which implements Eqs. (8), (11) and (13) directly.

4 Backgrounds and suppression, channel by channel

The rates of Sec. 3 say how many N's are produced; whether a channel is usable is decided by what else can imitate its two-vertex topology in 30 fb^{-1} of $b\bar{b}$ events. The organising fact, exploited throughout this section, is that *no single b-hadron SM decay produces a prompt $D\mu$ (or μ) system together with a second, downstream displaced vertex*: $b \to c$ is charged current, so one parent yields exactly one muon, and charm decays never produce an opposite-sign muon pair. Every background is therefore an *association of two unrelated objects* — a real or fake prompt system, plus a real or fake displaced vertex — and the strategy is uniform: purify the two vertices individually, then demand kinematic consistency between them.

For the decay side we use the branching ratios of the simulation chain's MadWidth table (Sec. 2.4), which agree with Ref. [6]: BR($N \to \mu q\bar{q}'$) $\simeq 50\%$, BR($N \to \mu\ell\nu$) $\simeq 17\%$, BR($N \to \nu q\bar{q}$) $\simeq 21\%$, BR($N \to \nu\mu^+\mu^-$) $\simeq 4.9\%$, BR($N \to \nu e^+e^-$) $\simeq 1\%$, remainder invisible. The targeted regime is the displaced one, $|U_{\mu N}|^2 \sim 10^{-8}$–$10^{-10}$ (Sec. 5), where the N vertex is measurably separated from the B vertex.

4.1 Universal handles

- **Two-vertex topology.** A displaced B vertex (flight $\mathcal{O}(1$ cm) at LHCb boosts, VELO vertex resolution $\mathcal{O}(10$–$100\ \mu$m)) containing the prompt system ($D\mu$, or μ alone), and an N vertex further downstream, inconsistent with both the PV and the B vertex in separation significance.

- **N-mass window and resonance vetoes.** The displaced-vertex mass ($m(\mu\mu)$ or $m(\mu h)$) must sit in a window around the tested m_N, excluding K_S (0.498), ϕ (1.020, BR($\phi \to \mu\mu$) $= 2.9 \times 10^{-4}$), the misID D^0 peak (just below 1.87), J/ψ (3.097) and $\psi(2S)$ (3.686) — noting that the J/ψ veto removes $m_N \simeq 2.9$–3.2 GeV whenever a D tag keeps that region kinematically open.

- **Visible/corrected mass below m_B.** With the flight direction as the momentum direction, $m_{\rm corr} \equiv \sqrt{m_{\rm vis}^2 + p_{T,\rm miss}^2} + p_{T,\rm miss} \le m_B$ for any signal combination; associations of particles from two different b hadrons largely overshoot. This is the single sharpest anti-combinatorial cut.

- **Same-sign muons (Majorana).** Requiring the prompt muon and a decay muon to have equal charge isolates the lepton-number-violating half of a Majorana N sample; genuine

SM sources of like-sign $\mu\mu$ at a common vertex are essentially absent, leaving only misID and random pairs. Cost: a factor two of signal (all of it for a Dirac N).

- **PID, isolation, and material vetoes.** Muon identification (hadron misID $\mathcal{O}(1\%)$ per track); both vertices isolated from other tracks (random-crossing vertices are not); $K_S \to \pi\pi$, $\Lambda \to p\pi$ and photon-conversion vetoes; displaced-vertex tracks re-tested under hadron mass hypotheses to reject mis-reconstructed $D \to K\pi(\pi)$ and Λ_c vertices.
- **Control samples.** Same-sign data, $m(\mu\mu)/m(\mu h)$ sidebands, fake-D mass sidebands, and the prompt (non-displaced) sample measure every background class from data; BR($B \to D^*\mu\nu$) provides the normalisation mode, as in Ref. [7].

4.2 $B \to D^{(*)}\mu N$ with $N \to \nu\mu^+\mu^-$: $D\mu$ tag + displaced dimuon

Signature. Prompt $D^{\pm}\mu^{\mp}$ (or D^* with the soft pion) at the B vertex; a displaced OS $\mu^+\mu^-$ vertex with $m(\mu\mu) = m_N$; invisible ν. Reach: $0.2 \lesssim m_N \lesssim 3.3$ GeV; $\mathrm{BR}_{\mathrm{prod}} \times \mathrm{BR}_{\mathrm{dec}} \sim 10^{-3}\,|U_{\mu N}|^2$ at $m_N = 1$ GeV (Table 2).

Backgrounds. (i) *Real $B \to D\mu\nu$ (*BR $\simeq 2.3\%$*, the exact prompt system) plus an accidental displaced OS pair* — cascade muons from the other b hadron forming a random crossing, pions decaying in flight (DIF), or $J/\psi \to \mu\mu$ from the other B. This is the dominant class. (ii) *Double-misID at a real displaced hadron vertex* ($D^0 \to K\pi$, $D^+ \to K\pi\pi$, $K_S \to \pi\pi$, charmless $B \to hh$): peaks in $m(\mu\mu)$ at the parent mass, dangerous in the $m_N \sim 1.87$ GeV window. (iii) *Fully combinatorial* real D + real prompt μ + real displaced $\mu\mu$ from different parents, including the $B_c \to J/\psi(\mu\mu)\mu\nu$ plus stray-D combination. (iv) *J/ψ from the other B*, peaking at $m(\mu\mu) = 3.097$ GeV, inside the reachable range with a D tag.

Suppression. The D–μ vertex fit removes classes (iii) and the stray-muon part of (i); $m(D\mu\mu\mu) < m_B$ (corrected mass) removes two-B associations; the N-vertex displacement (beyond the B-flight scale) removes J/ψ-from-B and DIF kinks; the $m(\mu\mu)$ vetoes of Sec. 4.1 remove (ii) and (iv); muon PID controls the residual misID. The final-state charge content ($\mu^{\mp}$ prompt + $\mu^+\mu^-$) always provides a like-sign $\mu\mu$ combination as a data-driven combinatorial control. Fake D's are sideband-subtracted; a D^* tag (soft-pion Δm cut) suppresses them further at some efficiency cost.

4.3 $B \to D^{(*)}\mu N$ with $N \to \mu q\bar{q}'$: $D\mu$ tag + displaced μ+ hadrons

Signature. As above, but the displaced vertex is $\mu^{\pm}\pi^{\mp}$ ($\mu\rho$, μ+jets at higher m_N) — the largest-rate channel, $\mathrm{BR}_{\mathrm{prod}} \times \mathrm{BR}_{\mathrm{dec}} \sim 10^{-2}\,|U_{\mu N}|^2$ at $m_N = 1$ GeV. Crucially, *nothing is invisible*: $m(D\mu\mu\pi) = m_B$, $m(\mu\pi) = m_N$, $m(K\pi\pi) = m_D$ — three simultaneous mass constraints. Reach: $m_N \lesssim 3.3$ GeV.

Backgrounds. (i) *Prompt physics background $B \to D^{**}\mu\nu$ and non-resonant $D\pi\mu\nu$* (BR $\sim$ few $\times 10^{-3}$): real D, μ, π from one B, but all at the B vertex and with non-peaking $m(\mu\pi)$ — irreducible only in the zero-displacement tail (large $|U_{\mu N}|^2$), where it is fitted via the $m(\mu\pi)$ shape and the missing mass (background has a ν, signal none). (ii) *Real $B \to D\mu\nu$ + fake displaced μh vertex*: misID at a real charm vertex ($D' \to K\pi\pi$, $\Lambda_c \to pK\pi$) or DIF + cascade muon; the hadron-hypothesis retest and PID remove most of it. (iii) *Combinatorial* associations as in Sec. 4.2, now suppressed by the triple mass constraint. (iv) Cross-feed *within* Table 2: $D^* \to D\pi^0/\gamma$ with the neutrals missed is indistinguishable from a D tag — the D and D^* production rows must be treated as one inclusive $D^{(*)}$ sample.

Suppression. The triple mass peak (B, N, D) with the missing-mass-zero requirement is close to hermetic; the same-sign variant (prompt μ^- with displaced $\mu^-\pi^+$) removes the SM cascade sources entirely and is the cleanest version of this channel. Residual background is estimated from the $m(\mu\pi)$ sidebands.

4.4 $B^+ \to \mu^+ N$ with $N \to \mu^\mp\pi^\pm$: fully reconstructed $B \to \pi\mu\mu$

Signature. The classic Majorana channel of Ref. [7]: no tag needed, $m(\mu\mu\pi) = m_B$ and $m(\mu\pi) = m_N$ with no missing energy; reach $0.3 \lesssim m_N \lesssim 5.2$ GeV, $\mathrm{BR}_{\mathrm{prod}} \times \mathrm{BR}_{\mathrm{dec}} \sim 5 \times 10^{-5}\,|U_{\mu N}|^2$ near the $m_N \simeq 3$ GeV production peak (Table 2).

Backgrounds (those of Ref. [7]): (i) *cascade $B \to D\mu\nu$ with $D \to \pi\mu\nu X$* — real $\mu\mu\pi$ but non-peaking in both masses, and dominant only in the OS sample; (ii) *misID*: $B \to D(\to K\pi\pi)\mu\nu$ with a hadron misidentified, and $B^+ \to \bar{D}^0(\to K^+\pi^-)\pi^+$ with double misID, peaking near $m(\mu\mu) \simeq m_D$; (iii) $B_c \to J/\psi X$ and J/ψ tails (veto); (iv) random combinations from $b\bar{b}$, dominant in the SS sample only at loose displacement; (v) the *irreducible SM decay* $B^+ \to \pi^+\mu^+\mu^-$ ($\mathrm{BR} \simeq 1.8\times10^{-8}$): same final state, but prompt at the B vertex — irrelevant for a displaced search, dominant for a prompt one.

Suppression. The double mass peak (B and N) plus the $J/\psi/\psi(2S)$ and D-mass vetoes; the same-sign selection ($\pi^\pm\mu^\pm\mu^\pm$) reduces all SM sources to misID and random pairs, which Ref. [7] showed are controllable to $\mathcal{O}(1)$ event; the displaced-N requirement removes (v) and the prompt cascade tail.

4.5 $B^+ \to \mu^+ N$ with $N \to \nu\mu^+\mu^-$: prompt muon + displaced dimuon

Signature. Three muons and missing energy, no hadronic tag; reach $0.2 \lesssim m_N \lesssim 5.2$ GeV, $\mathrm{BR}_{\mathrm{prod}} \times \mathrm{BR}_{\mathrm{dec}} \sim 10^{-5}\,|U_{\mu N}|^2$ at the production peak. This is the *only* light-B channel alive at $m_N \gtrsim 2.5$ GeV once the semileptonic modes close (Sec. 5).

Backgrounds. (i) *Prompt μ from a semileptonic B decay + displaced $\mu\mu$ from the other B* ($J/\psi \to \mu\mu$, cascade random crossing); (ii) $B_c \to J/\psi(\mu\mu)\mu\nu$ — a *real* three-muon final state at a real (mildly displaced) B_c vertex, peaking at $m(\mu\mu) = m_{J/\psi}$; (iii) DIF/misID combinations as before. With a ν lost, the only mass handles are $m(\mu\mu) = m_N$ and $m(\mu\mu\mu) < m_B$, so this channel is intrinsically more combinatorial than Secs. 4.2 and 4.4.

Suppression. J/ψ and $\psi(2S)$ vetoes (both reachable here), $m(\mu\mu\mu) < m_B$ corrected mass, N-vertex displacement beyond the B_c-flight scale for (ii), tight vertex isolation. The like-sign $\mu\mu$ combinations again provide the control sample.

4.6 B_c channels

$B_c \to J/\psi\,\mu N$ with $J/\psi \to \mu^+\mu^-$ and $N \to \nu\mu^+\mu^-$. A five-muon final state at three vertices (prompt $J/\psi\mu$ at the B_c point, displaced $\mu\mu$). Requiring $m(\mu\mu)_{\mathrm{tag}} = m_{J/\psi}$ plus the two-vertex topology leaves essentially only random associations in busy events — in practice a background-free counting experiment whose limitation is the parent rate ($f_c/f_u \simeq 3.9\times10^{-3}$), not background. Same for $N \to \mu q\bar{q}'$ ($J/\psi + 2\mu + \pi$).

$B_c^+ \to \mu^+ N$ (two-body). The tri-muon analysis of Sec. 4.5 (or $\mu\mu\pi$ of Sec. 4.4 with a softer vertex), extended to the unique reach $m_N < 6.2$ GeV; it is the *only* B-decay source at 5.2–6.2 GeV (Table 2), where the $W \to \mu N$ comparison channel is still rate-competitive.

4.7 Electron, tau and invisible N decays

- $N \to \nu e^+ e^-$ and $N \to \mu e \nu$ ($\sim 9\%$ combined): displaced vertices with electrons are VELO-only at LHCb; the dominant backgrounds are photon conversions (vetoed by mass and geometry), π^0 Dalitz decays and bremsstrahlung-degraded momentum. Useful as a cross-check of the muon channels rather than a primary search.
- $N \to \mu\tau\nu$ ($\sim 8\%$): a three-vertex chain (B; τ; $\tau \to \mu\nu\nu$ or $\tau \to 3\pi$) — virtually no SM source, striking topology, small rate.
- $N \to \nu q\bar{q}$ ($\sim 21\%$): displaced dijet vertex with the prompt $D\mu$ system; backgrounds are hadronic interactions in material and K_S/Λ decays rather than combinatorial muons. Statistically the largest untapped mode, experimentally the hardest; left for a dedicated VELO-material-veto study.

4.8 Summary

production	decay	m_N reach	$\mathrm{BR}_{\mathrm{prod}} \times \mathrm{BR}_{\mathrm{dec}}/\lvert U_{\mu N}\rvert^2$	dominant background	decisive handle
$B \to D^* \mu N$	$\nu\mu\mu$	$\lesssim 3.3$	$\sim 10^{-3}$	real $B \to D\mu\nu$ + accidental $\mu\mu$	D–μ vertex, $m_{\mathrm{corr}} < m_B$
$B \to D^* \mu N$	$\mu q\bar{q}'$	$\lesssim 3.3$	$\sim 10^{-2}$	$B \to D^{**}\mu\nu$ (prompt physics)	triple mass peak, missing mass, SS
$B^+ \to \mu N$	$\mu q\bar{q}'$	$\lesssim 5.2$	$\sim 5 \times 10^{-5}$	cascade $B \to D\mu\nu$, misID	B+N double peak, SS [7]
$B^+ \to \mu N$	$\nu\mu\mu$	$\lesssim 5.2$	$\sim 10^{-5}$	prompt μ + other-B J/ψ	J/ψ veto, $m(3\mu) < m_B$
$B_c \to J/\psi \mu N$	$\nu\mu\mu$	$\lesssim 3.2$	$\sim 3 \times 10^{-4}$ ($\times f_c$)	combinatorial only	5-muon, 3-vertex topology
$B_c^+ \to \mu N$	$\nu\mu\mu/\mu q\bar{q}'$	$\lesssim 6.2$	$\sim 10^{-5}$–10^{-2} ($\times f_c$)	as tri-muon channel	only source > 5.2 GeV

Table 4: Background structure of the major B-production $\times$ N-decay signatures (Sec. 4.2–4.7). Rates are the peak values of Table 2 times the decay BRs of Sec. 2.4; "$\times f_c$" recalls the B_c parent suppression of Table 1. SS = same-sign muon selection (Majorana). All channels share the universal handles of Sec. 4.1.

The pattern is uniform: wherever a full or near-full kinematic constraint exists (D tag, B-mass edge, double mass peak), the irreducible backgrounds are prompt physics processes that die under the displacement requirement, and the displaced regime is combinatorial-limited; wherever the constraint is lost (tri-muon, $\nu q\bar{q}$), the search is combinatorial-limited already at moderate displacement. The $D^{(*)}$-tagged channels therefore offer the best sensitivity per produced N at $m_N \lesssim 3$ GeV, while the leptonic $B^+ \to \mu N$ and $B_c^+ \to \mu N$ modes extend coverage to $m_N \simeq 5.2$ and 6.2 GeV respectively at the price of looser kinematic closure.

5 Discussion

5.1 Channel hierarchy as a function of mass

The mass dependence in Tables 2 and 3 partitions the m_N range into distinct regimes:

- $m_N \lesssim 2$ GeV: **semileptonic dominance**. $B \to D^* \mu N$ and $B \to D\mu N$ (plus $\Lambda_b \to \Lambda_c \mu N$) produce $\mathcal{O}(10^{10})$ HNLs per unit $\lvert U_{\mu N}\rvert^2$ in the LHCb acceptance — about three orders of magnitude above $W \to \mu N$. The leptonic channel is a percent-level correction here.
- $2 \lesssim m_N \lesssim 3.3$ GeV: **crossover**. The semileptonic retention factor collapses ($R \sim$ few% at 2 GeV, $\sim 10^{-4}$–10^{-3} at 3 GeV) while the helicity-enhanced leptonic rate rises to its

$\sim 2 \times 10^{-4} \, |U_{\mu N}|^2$ plateau. From $m_N \simeq 2.5$ GeV onward $B^+ \to \mu^+ N$ is the leading light-B channel.

- $3.3 \lesssim m_N \lesssim 5.2$ GeV: **leptonic only**. $B^+ \to \mu^+ N$ is the sole significant channel (the $b \to u$ semileptonic tail is 2–3 orders down), still producing ∼3–30× more HNLs than the W channel per unit $|U_{\mu N}|^2$.
- $5.2 \lesssim m_N \lesssim 6.2$ GeV: B_c **niche**. $B_c^+ \to \mu^+ N$ alone is open, at $\sim 10^8$–10^9 per unit $|U_{\mu N}|^2$.
- $m_N \gtrsim 6.2$ GeV: W **territory** — the regime of the existing CMS/LHCb sensitivity chain.

The per-electron-mixing numbers are identical at $m_N \gg m_\mu$ (the helicity enhancement is flavour-blind), so the same tables apply to $|U_{eN}|^2$ with the obvious final-state relabelling; τ-coupled HNLs are produced via $B \to X\tau N$ and $B^+ \to \tau^+ N$ for $m_N < 3.5$ GeV.

5.2 Production is not the bottleneck — lifetime and trigger are

All rates above scale $\propto |U_{\mu N}|^2$, while the lab-frame flight distance scales $\propto \beta\gamma/|U_{\mu N}|^2$. At $|U_{\mu N}|^2 = 1$ the HNL is prompt ($c\tau \sim 10^{-7}$ mm at $m_N = 2$ GeV); a displaced-decay search in the LHCb muon-station window ($z \simeq 11.5$–18 m) requires $\beta\gamma c\tau \sim 10$ m, i.e. $|U_{\mu N}|^2 \sim 10^{-9}$–$10^{-10}$ at these masses. Rescaling Table 3: even the dominant $B \to D^* \mu N$ channel then yields only $\mathcal{O}(10^2)$ *produced* HNLs at $m_N = 1$ GeV, before the decay-in-window probability (itself a strong function of the HNL boost spectrum from B decays, which is much softer than from W decays). The enormous production advantage of the B sector is therefore exactly what buys access to the long-lifetime, small-$|U_{\mu N}|^2$ regime — and the quantitative question is the convolution of the B-decay HNL boost spectrum with the decay-in-window probability. That convolution is the natural next step (a Pythia $b\bar{b}$ sample with forced $B \to X\mu N$ decays, through the same muon-detector efficiency machinery as the W chain).

The second structural difference is the trigger: $W \to \mu N$ provides a hard, isolated, prompt muon ($p_T \sim 40$ GeV) that makes the existing samples nearly background-free, while HNLs from B decays come with soft muons in busy $b\bar{b}$ events. A B-production analysis must replace that with a displaced-vertex selection (VELO, $m_N \gtrsim$ few GeV phase space allowing) or a muon-station-shower self-trigger. The channel-by-channel background structure and the corresponding handles are the subject of Sec. 4.

5.3 Caveats on the numbers

- The semileptonic retention factors use the $\lambda^{3/2}$ constant-form -factor spectrum model; realistic form factors (CLN/CLN-like for $B \to D^{(*)}$, LCSR for $B \to \pi$) shift $R(m_N)$ by ∼30–50%. The leptonic two-body rates, being pure phase space × decay constant, are exact up to the inputs f_B, f_{B_c} and $\mathrm{BR}(B_c \to \mu\nu)$ (∼25% theory).
- Fragmentation fractions are p_T-dependent (notably f_{Λ_b}), and f_c carries its own ∼25% uncertainty.
- The $b\bar{b}$ cross section is the FONLL/LHCb-measured value; per-species cross sections inherit its ∼10% systematic.
- Table 3 counts *produced* HNLs in $2 < \eta < 5$ for the B sector but 4π for the W row; the in-acceptance W number is ∼15–20% of the quoted value.
- Angularity: nothing here yet accounts for where the HNL flies relative to the LHCb arm — the B-decay HNL boost is $\mathcal{O}(\text{few GeV}/m_N)$, so the fraction of produced HNLs pointing into $2 < \eta < 5$ with the right $\beta\gamma$ is the subject of the follow-up study mentioned above.

References

[1] LHCb Collaboration, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, arXiv:1612.05140 (erratum-corrected $\sigma(b\bar{b}, 2 < \eta < 5) = 154.3 \pm 1.5 \pm 14.3$ μb at 13 TeV).

[2] LHCb Collaboration, *Measurement of b-hadron fractions in 13 TeV pp collisions*, Phys. Rev. D **100** (2019) 031102(R), arXiv:1902.06794.

[3] LHCb Collaboration, *Measurement of B_c^- production in proton-proton collisions at $\sqrt{s} = 7$ and 13 TeV*, arXiv:1910.13404.

[4] ATLAS Collaboration, *Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **759** (2016) 601, arXiv:1603.09222; NNLO reference cross sections (FEWZ/DYNNLO).

[5] Particle Data Group, *Review of Particle Physics*, PTEP **2022** (2022) 083C01 and 2024 update.

[6] K. Bondarenko, A. Boyarsky, D. Gorbunov, O. Ruchayskiy, *Phenomenology of GeV-scale heavy neutral leptons*, JHEP **11** (2018) 032, arXiv:1805.08567.

[7] LHCb Collaboration, *Search for Majorana neutrinos in $B^- \to \pi^+\mu^-\mu^-$ decays*, Phys. Rev. Lett. **112** (2014) 131802, arXiv:1401.5361.