Chapter 9

Selection of Heavy Neutral Leptons

The selection of Heavy Neutral Lepton (HNL) signals from Standard Model (SM) neutrino and cosmic backgrounds using Monte Carlo (MC) is presented. MC samples were simulated using the framework detailed in Chapter 5 and reconstructed using the framework detailed in Chapter 6. This is the very first exploration of Short-Baseline Near Detector (SBND) physics capabilities to search for HNLs, and this selection provides the first benchmark for understanding the current reconstruction performance of HNL signals based on MC. The selection exploits the boosted topology and late arrival features of HNL signals using the reconstructed charge and light signals from the Time Projection Chamber (TPC) and Photon Detection System (PDS) combined. This set up the ground work that can be carried out on data once the detector is operational.

The following chapter covers details on the selection workflow to identify HNL signals from backgrounds. The foundation of the selection is given in Section 9.1, including a description of signals and backgrounds, the generated MC samples and relevant parameter and distributions to evaluate to selection. The first stage of the selection is to reject cosmic backgrounds as discussed in Section 9.2 and the second stage is to reject SM neutrino backgrounds as discussed in Section 9.3. Then, Section 9.4 contains details the last stage of the selection to identify HNL showers from shower-like backgrounds. The result of the selection procedure is summarised in Section 9.5. A discussion of possible improvements in sensitivity with better timing resolution is detailed in Section 9.6. Finally, some concluding remarks are provided Section 9.7.

9.1   Selection Introduction

This section provides details of all the ground work before performing the selection. Definitions of signals are backgrounds are presented in Section 9.1.1. Descriptions of MC samples used in the selection are provided in Section 9.1.2. Parameters to evaluate the selection, including definitions of efficiency and the beam bucket distribution, are detailed in Sections 9.1.3 and 9.1.4 respectively.

9.1.1   Signal and Background Definition

The selection begins with defining the signal topology, namely $\pi^0 \to \gamma\gamma$ showers resulting from HNLs decaying inside the Fiducial Volume (FV) of the SBND detector. FV is a smaller volume approximately 70% of the active volume, to be defined in the forthcoming Section 9.3.1. The di-photon showers of HNLs result in one or more showers without any hadronic activities at the vertex. Fig. 9.1a shows an event display of two separable photon showers, where each shower can be seen distinctively. In the case where only a single shower is reconstructed, two scenarios can happen. The first scenario is that only a single photon shower deposits energy inside the detector while the other one escapes. The second scenario is that the di-photon showers are very boosted and forward-going. Fig. 5.4 in Section 5.2 shows that the angle of π^0 to the beam direction is very small $< 20°$ for HNLs in the mass range of 140-260 MeV. Thus, the resulting di-photon showers can overlap each other, in which case the opening angle between the two showers is too small to be reconstructed as two distinct showers. Fig. 9.1b shows an event display of very boosted di-photons showers, which are likely to be reconstructed as a single energetic shower.

Given this signal topology, the first-order background topology from SM neutrinos is Neutral Current interactions that produce π^0 (NC π^0). This interaction type also produces di-photon showers with little or no hadronic activities at the vertex. The second-order background topology is from Charged Current electron (anti-)neutrinos (CC ν_e) interactions. This interaction type typically produces one or multiple hadrons in addition to a single shower. However, in some scenarios, the hadrons are too low in energy to be reconstructed, resulting in a single shower topology after reconstruction. Fig. 9.2 shows an event display of the observable di-photon showers from NC π^0 interaction, which is indistinguishable from the di-photon showers from HNLs. The key distinction separating HNL showers from these SM neutrino showers is the boosted topology of HNL showers, where HNL di-photon showers have smaller opening angles and tend to travel preferably in the beam direction.
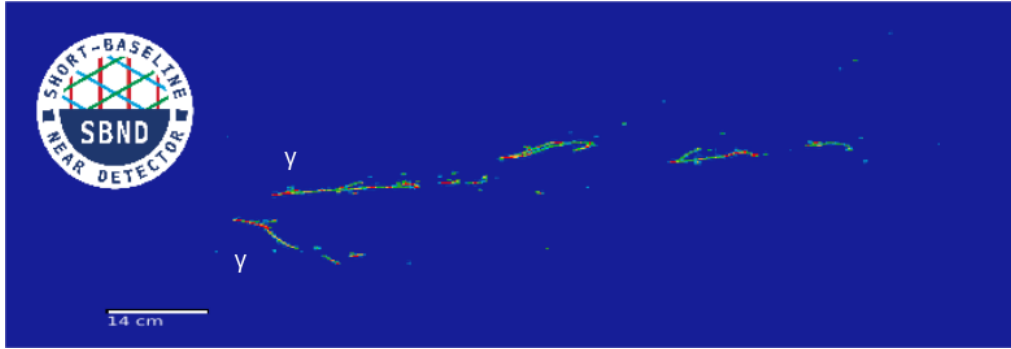
(a) Separable di-photon showers
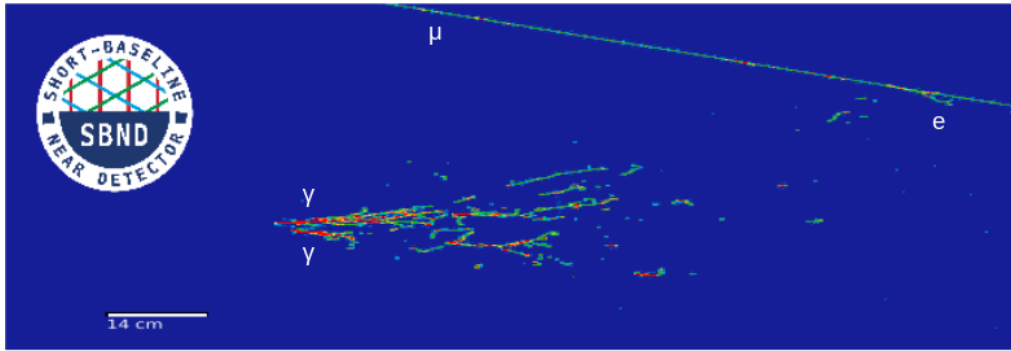
(b) Overlapped di-photon showers

Fig. 9.1 Event displays showing two common topologies of simulated di-photon showers from HNLs.
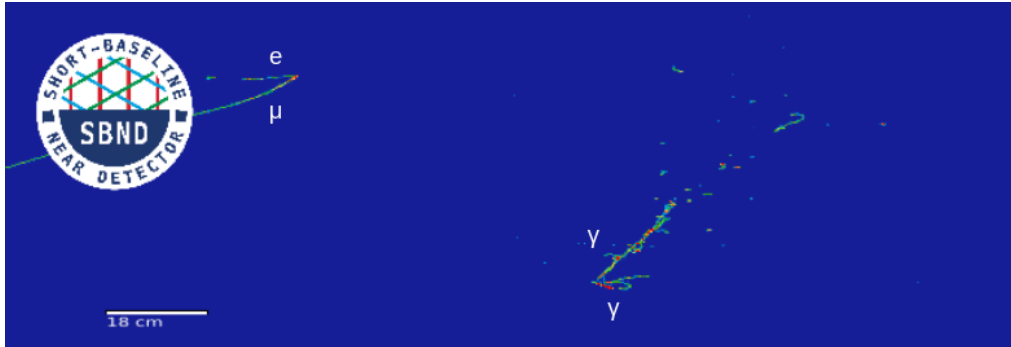

Fig. 9.2 Event display showing di-photon showers from a simulated NC π^0 interaction and a cosmic stopping muon.

SM neutrino interactions can occur outside the FV, but their products can have sufficient energy to propagate inside the FV. Interactions occurring outside the FV but inside the detector volume are referred to as Non-FV interactions. Interactions occurring completely outside the detector volume are referred to as dirt neutrino interactions. As previously discussed in Section 5.3.1, despite interacting outside of the FV, these interactions can

123

introduce non-negligible backgrounds, especially if their products also contain showers in the final states.

Finally, any background interactions that produce tracks are considered low-priority backgrounds since a track topology is easily distinguishable from a shower topology. From SM neutrinos, these interactions are from Charged Current muon (anti-)neutrinos (CC ν_μ) or any Neutral Current interactions that do not produce a neutral pion (Other NC). Fig. 9.3 shows an event display of a common observable from CC ν_μ interactions containing 1 muon and 1 proton in the final state. Similarly, cosmic muons typically leave very long tracks crossing the entire detector with features of delta rays or Michel electrons (See Sections 8.2 and 3.2.1). Fig. 9.1b (top right) and Fig. 9.3 (bottom left) both show a long cosmic track with some delta rays along the track. Fig. 9.2 (top left) shows a cosmic muon coming to a stop and decaying into a Michel electron.
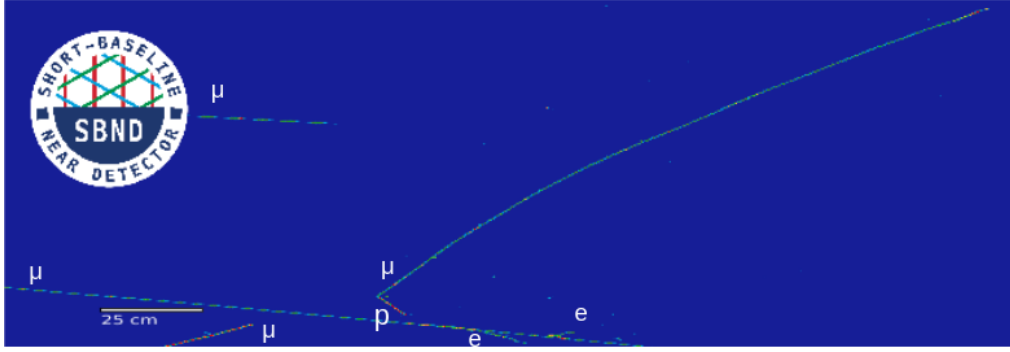

Fig. 9.3 Event display showing a muon and a proton track from a simulated CC ν_μ interaction together with a few cosmic muons.

9.1.2 Description of MC Samples

HNL signals were overlaid with cosmic muons occurring within the TPC readout window. Samples at the HNL mass of 140, 160, 180, 200, 220, 240 and 260 MeV were generated, totalling 7 samples with 60,000 events per sample. The number of events per sample can be re-weighted from the simulated coupling $|U_{\mu 4}|^2$ to another coupling $|U'_{\mu 4}|^2$ by applying a weight as follows:

$$ w = \left(\frac{|U_{\mu 4}|^2}{|U'_{\mu 4}|^2} \right)^2 . \tag{9.1} $$

Eq. 9.1 allows for the signal scaling required to perform the limits setting in Chapter 10.

Three samples of SM neutrinos were also generated. The first one is a core sample with all SM neutrino interactions occurring inside the detector volume as well as outside the

detector in the *Rockbox* volume, as discussed in Section 5.3.1. Two additional dedicated samples of enriched NC π^0 and CC ν_e backgrounds were also generated, to improve the limited statistics of these interactions in the core sample. The three samples were normalised to an exposure of 1×10^{21} Protons On Target (POT) to account for 3 years of data taking. This yields $\sim 331,000$ NC π^0 interactions and $\sim 33,000$ CC ν_e interactions which are the primary background. Other background from CC ν_μ and Other NC interactions make a total of ~ 5 million interactions. An additional ~ 2 million and ~ 3 million interactions from Non-FV and dirt interactions are also considered as backgrounds, although only a fraction of them deposit energy in the detector.

Finally, a cosmic-only sample was generated to account for in-time cosmics (See Section 5.3.2). This sample consists of events triggered by cosmic-only interactions. However, it is important to note that a dedicated trigger efficiency study will be carried out to better understand the rate of in-time cosmic events once SBND is operational. The cosmic-only sample was also normalised to the target POT, and combined with SM neutrino samples to form a single sample describing the background to the HNL signal.

The unit of the selection relies on *events*, where a single event corresponds to a single trigger (See Fig. 4.8 Section 4.2.7). After reconstruction, each event contains multiple *slices*, which are reconstruction units created by Pandora to describe interactions (See Section 6.2). A slice encapsulates all energy in the TPC from a single origin and consists of a hierarchy of particles starting from the interaction vertex, where each particle can resemble a track or a shower. The equivalent reconstruction unit to a slice from the PDS reconstruction is a *flash*, where a flash contains all the light produced from an interaction (See Section 6.3). The selection is performed on slices, where slices are accepted or rejected based on the series of cuts using the reconstructed information of the slice or by matching a slice to a flash.

9.1.3   Selection Efficiency Definition

For monitoring and quantifying the impacts of selection cuts, selection efficiencies are defined for signals and backgrounds respectively. The selection efficiency is defined as:

$$\text{Signal Efficiency} = \frac{\text{Number of selected signal slices with completeness } > 50\ \%}{\text{Number of signal slices reconstructed as neutrinos}}. \tag{9.2}$$

In the numerator, the requirement of > 50% completeness implies that at least 50% of the slice energy must be deposited by a HNL. This prevents double counting, such that only well-reconstructed signal slices are considered. In the denominator, the requirement for

slices to be reconstructed as neutrinos is embedded by the Pandora workflow. As stated in Section 6.2.3, Pandora performs a cosmic rejection very early in the reconstruction so that only neutrino-like slices are reconstructed. The Pandora cosmic rejection is also employed as the first cosmic cut detailed in Section 9.2.1. The denominator describes the starting number of signal slides and the numerator describes the selected signal slices. Thus, Eq. 9.2 describes the efficiency of signals remained after selection.

On the other hand, the background efficiency is defined as:

$$\text{Background Efficiency} = \frac{\text{Number of selected background slices}}{\text{Number of background slices reconstructed as neutrinos}}. \quad (9.3)$$

Here, the background slides encapsulates both SM neutrino and cosmic slices, that are reconstructed by Pandora as neutrino-like slices. The denominator describes the starting number of background slides and the numerator describes number of selected background slices. Eq. 9.3 therefore describe the efficiency of background slices being selected as signals.

The selection aims for a high background *rejection* without compromising the signal *efficiency*. This is equivalent to achieving a low background efficiency defined in Eq. 9.3 and a high signal efficiency defined in Eq. 9.2. Both of these efficiency numbers are discussed for each cut and included in the legends of the upcoming plots.

9.1.4 The Beam Bucket Distribution

The selection workflow was developed by exploiting distinct features of HNLs. One previously stated feature is the boosted topology of HNL showers as discussed in Section 9.1.1. Another feature is the late arrival of HNLs relative to SM neutrinos, as previously depicted in Fig. 5.7 in Section 5.2 showing the arrival time distribution of HNLs and SM neutrinos. The distribution of SM neutrinos resembles a Gaussian-shaped bucket as they travel nearly at the speed of light, whilst HNLs travel at a slower velocity and smear the Gaussian. It is referred to as the *beam bucket distribution* in this work. It is also the key distribution for setting the upper limits on the coupling $|U_{\mu 4}|^2$ of HNLs since it demonstrates the distinct shape difference between the signal and the background, which is required by the setting limits procedure to be discussed in Chapter 10.

To reconstruct the beam bucket distribution, the required information is the flash time matched to a slice that corresponds to the start time t_0 of the interaction, of which the timing reconstruction is detailed in Section 6.3.1. From the interaction time t_0, the arrival time at the upstream wall of the detector was computed by shifting from the interaction vertex

z-position to $z = 0$. The arrival time corresponds to 81 beam buckets in a single beam spill and thus, to overlay 81 buckets as a single one, a modulus equal to the spacing between buckets is applied. The spacing was measured to be 18.936 by the MicroBooNE experiment [1]. Discussion on different smearing contributors to the beam bucket reconstruction is given in Section 9.6. The beam bucket distribution are shown throughout this chapter to demonstrate the impacts of the selection.

9.2 Cosmic Background Removal

Cosmic rejection is the first step of selection, targeting two cosmic components: (1) in-time cosmics occurring inside the beam spill and (2) out-of-time cosmics occurring outside the beam spill but inside the readout window (See Section 5.3.2). The cosmic removal by Pandora is the first cut, presented in Section 9.2.1. The beam spill cut is given in Section 9.2.2 and the last cut employing a Boosted Decision Tree (BDT) is provided in Section 9.2.3.

9.2.1 Pandora Unambiguous Cosmic Removal

Being a surface detector, SBND is exposed to a high rate of cosmic rays, expecting ~ 185 million reconstructed slices from cosmics for the POT exposure of 1×10^{21}. As a comparison, the expected rate of reconstructed slices from SM neutrino interactions is ~ 11 million slices. The first cosmic rejection step targets primarily at removing out-of-time cosmic muons. As described in Sections 6.2.3 and 9.1.3, Pandora performs an unambiguous cosmic removal early in the reconstruction chain and only neutrino-like slices are reconstructed. The selection thus begins with selecting only slices reconstructed as a neutrino. This rejects 90% of the ~ 185 million slices from cosmic, leaving behind only 19.5 million slices. Meanwhile, only 0.6% of the reconstructed slices from HNL signals are removed, with similar reductions across different SM neutrino interactions.

9.2.2 Beam Spill Cut

The second cut to remove cosmics is to consider the flash time of a slice, corresponding to the start time t_0 of an interaction. Only slices matched to a valid flash are selected, where the slice-to-flash matching relies on the level of agreement between reconstructed energy from charge and light (See Section 6.4.1). Moreover, the time of the matched flash is required to be within the beam spill window. In the simulation of MC samples, the beam spill window

127

is configured to be between [0.367, 1.967] μs, with $t = 0$ μs corresponding to the first POT of a beam spill. Moreover, an interaction can occur anywhere along the 500 cm z-length of the detector, equivalent to a smearing of 17 ns in timing. Thus, the beam spill acceptance window is widened to [0.350, 1.984] μs. The beam spill is illustrated in Fig. 9.4a, with the acceptance window shown as red lines.

The cut rejects 4 million cosmic slices while minimally reducing signal efficiency by 3%. Fig. 9.4b shows the beam bucket distribution after applying the cut, where two components of cosmic rays can be observed. There is a flat distribution coming from out-of-time cosmics and a very small Gaussian-shaped distribution coming from in-time cosmics.

9.2.3   CRUMBS Cut

The third cut targets the out-of-time cosmic components by employing the CRUMBS score of a slice, which is scored by a BDT to distinguish between a neutrino-like slice and a cosmic-like slice (See Section 6.4.2). The score distribution of CRUMBS is plotted in Fig. 9.4c, showing a good separation between neutrino-like and cosmic-like. A cut is placed to reject any slices with CRUMBS scores less than 0, effectively removing 14 million of the remaining cosmic slices.

Comparison the beam bucket distribution before and after the CRUMBS cut, Fig. 9.4b and 9.4d, demonstrates that the majority of the removed cosmics are the out-of-time component. The remaining cosmic slices are the in-time component, concentrating at the centre of the beam bucket. This cut results in an effective background rejection as the background efficiency reduces more than half from 8.3×10^{-1} to 3.0×10^{-1}, whilst the signal efficiency only drops by 5%. By the end of the cosmic rejection, only $\sim 432,000$ of the starting 185 million cosmic slices remain, equivalent to a 99.9% removal of the cosmic background alone.

9.3   Neutrino Background Removal

The next set made up of three cuts focusing on rejecting backgrounds of SM neutrinos. The cut on detector volume is presented Section 9.3.1. The cut on the reconstruction quality is detailed in Section 9.3.2. Finally, the cut targeting at removing track-like particles from SM neutrino interactions is provided in Section 9.3.3.

(a) Beam spill cut

(b) After beam spill cut
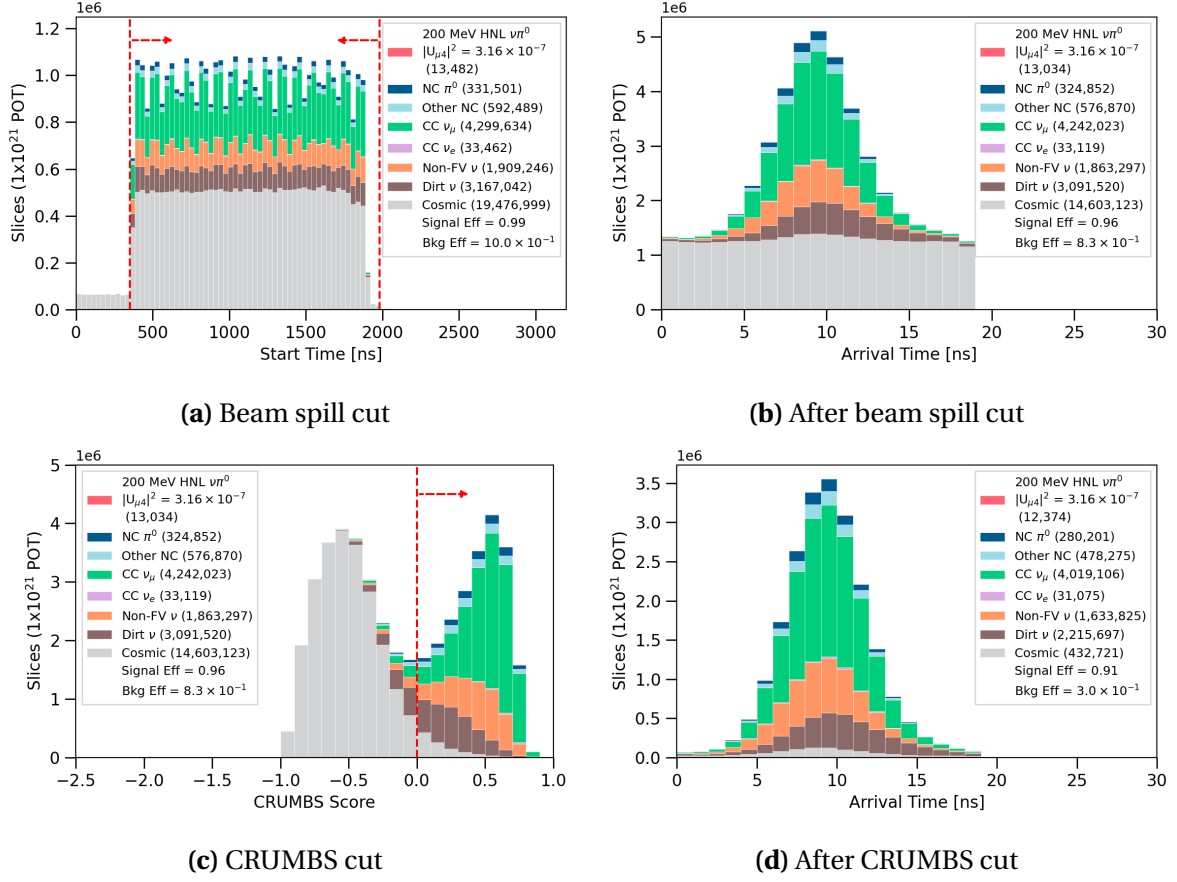
(c) CRUMBS cut

(d) After CRUMBS cut

Fig. 9.4 Cosmic rejection cuts (left) and the beam bucket distribution after each cut (right).

9.3.1 Fiducial Volume Cut

The cut on detector volume aims to remove backgrounds from Non-FV neutrinos and dirt neutrinos that interact outside of the FV but for which their products can deposit energy inside the FV. The cut requires the reconstructed vertex of a slice to be inside the FV, which is approximately 70% of the entire active volume of the detector. The FV is defined as follows:

- x-position: $-180 < x < -5, 5 < x < 180$ cm,

- y-position: $-180 < y < 180$ cm,

- z-position: $10 < z < 450$ cm.

The boundary is set on the x-axis to reject vertices reconstructed close to the anode and cathode. Vertices close to the cathode means the charge clusters must traverse the full drift

distance before reaching the anode for detection, therefore, are more susceptible to detector effects (See Section 3.3.1), resulting in poor reconstruction. Meanwhile, vertices close to the anode might also indicate particles entering from the side of the detector which are likely to be cosmic muons and Non-FV/dirt neutrino backgrounds. The boundary on the y-axis rejects interactions that might enter the detector from the top, such as cosmic muons, or bottom, such as Non-FV/dirt neutrinos. Finally, the boundary on the z-axis for $z > 10$ cm rejects entering particles and $z < 450$ cm requires enough downstream volume for a shower to grow. Overall, these cuts ensure the quality of reconstruction.

The distribution of vertices reconstructed inside and outside of the FV is shown in Fig. 9.5a and the result of the cut is demonstrated in Fig. 9.5b. Dirt neutrino slices reduce from ~ 2 million slices to only ~306,000 slices while Non-FV neutrino slices drop from ~1.6 million slices to only ~ 99,000 slices. The cut reduces both the background efficiency and signal efficiency by a third as it is consistent with rejecting 30% of the detector volume.

9.3.2 Number of Hits Cut

This cut aims to select well-reconstructed slices by examining the number of hits of the primary particle in a slice that deposits the most energy. The number of hits is particularly important given that Pandora relies on hit information to reconstruct 3D information of particles in a slice (See Sections ?? and 6.2.3). The more hits associated with a particle, the more information is available for Pandora to reconstruct its topology and calorimetry. The number of hits requirement for the primary particle is ≥ 50 hits to provide sufficient information for a reliable Pandora reconstruction. Fig. 9.5c demonstrates the distribution of the number of hits of the primary particle in a slice. Only the first bin is rejected by this cut, demonstrating that only a small amount of slices containing primary particles with < 50 hits, which are likely to be poorly reconstructed. The beam bucket distribution after the cut is shown in Fig. 9.5d, where it can be seen that the cut reduces the signal and background efficiency minimally by < 1%.

9.3.3 SM Neutrino Track Removal

The next sets of cuts focus on rejecting SM neutrino backgrounds that produce tracks originating from muons, protons and charged pions. The cut uses the score distribution from the Razzled BDT (See Section 6.4.3). There are two types of Razzled variables examined for this cut: (1) the number of p, μ, π in a slice as identified by Razzled and (2) the Razzled

(a) Fiducial volume cut

(b) After fiducial volume cut

(c) Number of hits cut

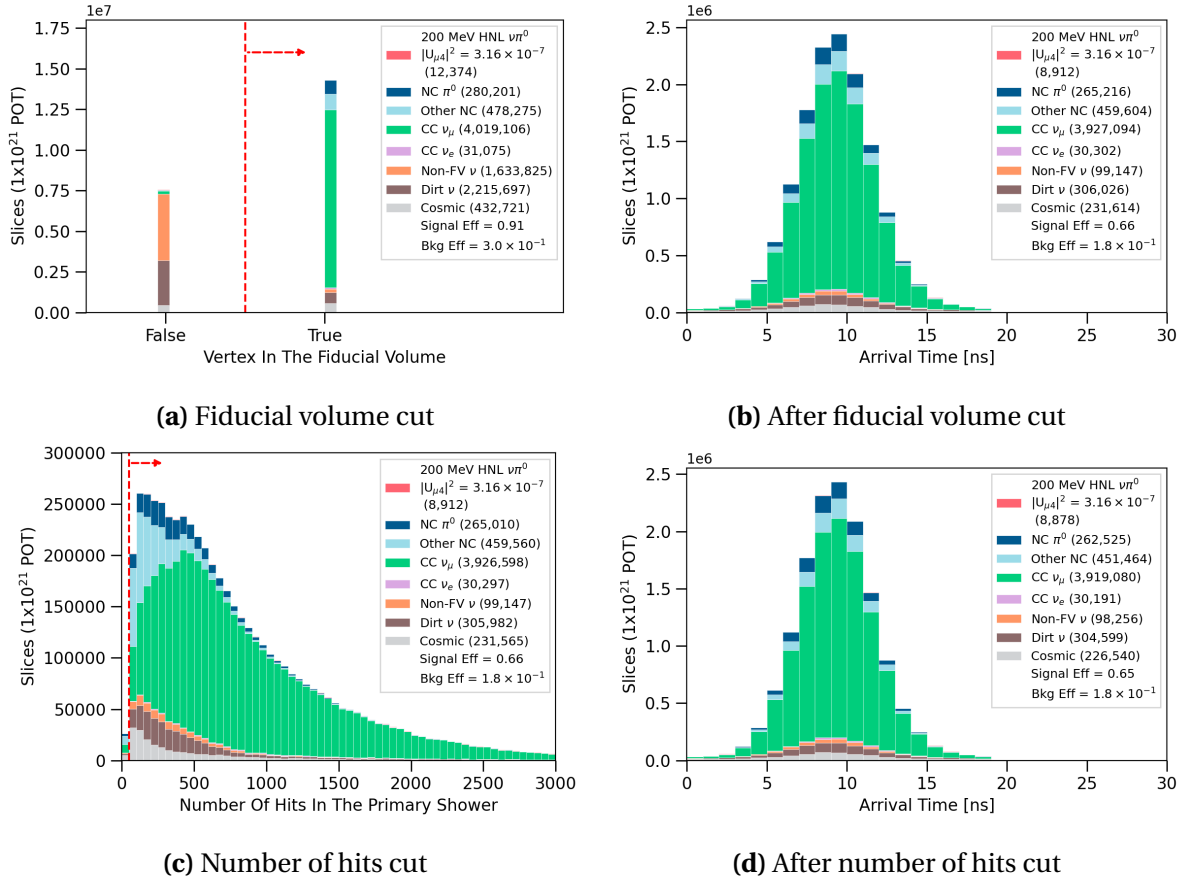**(d)** After number of hits cut

Fig. 9.5 Fiducial volume and number of hits cuts (left) and the beam bucket distribution after each cut (right).

p, μ, π scores of all particles in a slice. The former cut relies on Razzled assigning a type to a particle based on its highest particle type score from the BDT. The latter cut is to further reject slices if they contain particles with a Razzled score higher than a chosen threshold.

Fig. 9.6a and 9.6b demonstrate the two cuts respectively for rejecting muon-like particles. Fig. 9.6a shows the requirement on the number of Razzled-identified muons is 0 while Fig. 9.6b shows that only slices containing particles with Razzled muon score < 0.04 are selected. The cuts are very aggressive without compromising signal efficiency due to the distinction between HNL signals and muon tracks. Comparison between the beam bucket distribution before and after the muon cut, Fig. 9.5d and Fig. 9.6c, can be made to evaluate the impacts of the cut. The muon cut effectively rejects 96% of the 4 million CC ν_μ slices, leaving only $\sim 161,000$ slices remaining. HNL slices are also affected by the cut such that the signal efficiency reduces from 65% to 51%.

Similar cuts are then applied consecutively to reject protons and charged pions. As with the muon cuts, the cuts on the number of p, μ, π as identified by Razzled in a slice are very aggressive to require that the selected slices not contain any track-like particles. On the other hand, the cuts on the Razzled p, μ, π scores were optimised for each particle type to maximise the background rejection without costing the signal efficiency. Additional conditions are required on the reconstructed Kinetic Energy (KE) to be > 32.7 MeV for protons and < 32.1 MeV for charged pions to ensure particles are well-reconstructed. To summarise, the cuts to reject muons, protons and pions are as follows:
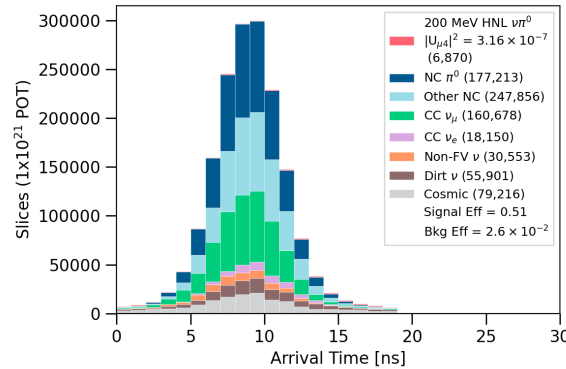
1. Muon cuts:

   • Number of Razzled-identified muons in a slice = 0,

   • Slices containing only particles with Razzled muon score < 0.04.

2. Proton cuts:
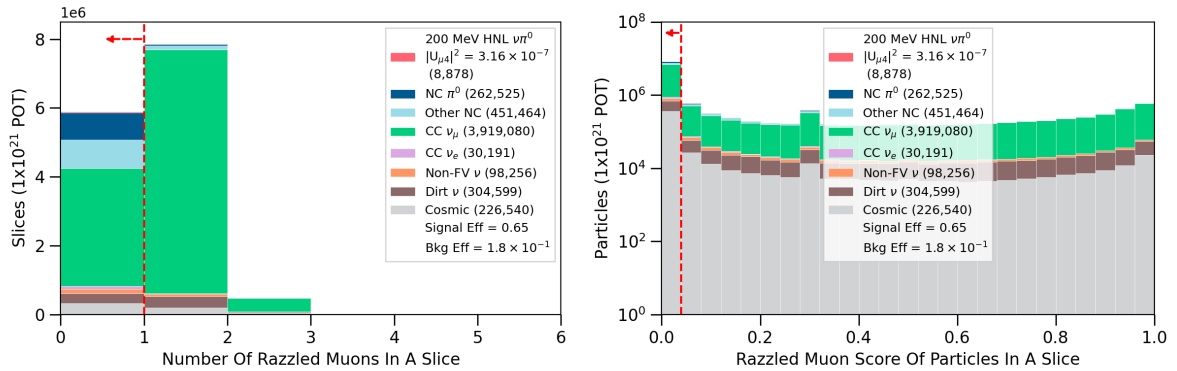
(a) Number of Razzled-identified muons cut

(b) Particles with Razzled muon score cut

(c) After muon cut

Fig. 9.6 Muon cuts (top) and the beam bucket distribution after the cuts (bottom).

- Number of Razzled-identified protons with KE > 32.7 MeV in a slice = 0,
- Slices containing only particles with Razzled proton score < 0.96.

3. Pion cuts:

- Number of Razzled-identified pions KE > 32.1 MeV in a slice = 0,
- Slices containing only particles with Razzled pion score < 0.82.

The cuts to reject protons are illustrated in Fig. 9.7a and 9.7b. The impacts of the proton cut can be seen in the beam bucket distribution in Fig. 9.7e, as any interactions producing protons are removed, significantly reducing SM neutrino backgrounds. The most impacted interaction modes are Other NC interactions reducing from $\sim 249,000$ to $\sim 17,000$ slices, CC ν_μ interactions reducing from $\sim 161,000$ to $\sim 46,000$ slices and NC π^0 interactions reducing from $\sim 177,000$ to $\sim 88,000$ slices.

The pion cuts are depicted in Fig. 9.7c and 9.7b. The result of the pion cut can be observed in the beam bucket distribution shown in Fig. 9.7f, where the cut further cleans up any SM neutrino slices that are not already rejected by the muon and proton cuts. CC ν_μ interactions are the most affected, decreasing from $\sim 46,000$ to $\sim 30,000$ slices. This is followed by a reduction of Other NC interactions from $\sim 17,000$ to $\sim 9,000$ slices.

The background efficiency at the end of the track removal significantly decreases by two orders of magnitudes from $\mathcal{O}(10^{-1})$ to $\mathcal{O}(10^{-3})$, demonstrating the effectiveness of these cuts. Meanwhile, the HNL signal efficiency only decreases by 65% to 46%.

9.4 HNL Shower Selection

After the track removal, the next five cuts targets at identifying HNL showers from shower-like backgrounds. The electron shower cut is provided in Section 9.4.1. The track score cut, to keep only very shower-like signals, is detailed in Section 9.4.2. In following, the calorimetry and theta cuts in Sections 9.4.3 and 9.4.4 exploit the boosted topology of HNL showers. Finally, since the π^0 invariant mass can be reconstructed, it is also used as a cut variable to be discussed in Section 9.4.5.

9.4.1 Electron Shower Removal

The first cut of this HNL shower selection aims at rejecting showers originating from electrons. The key differences between electron showers and photon showers are the shower
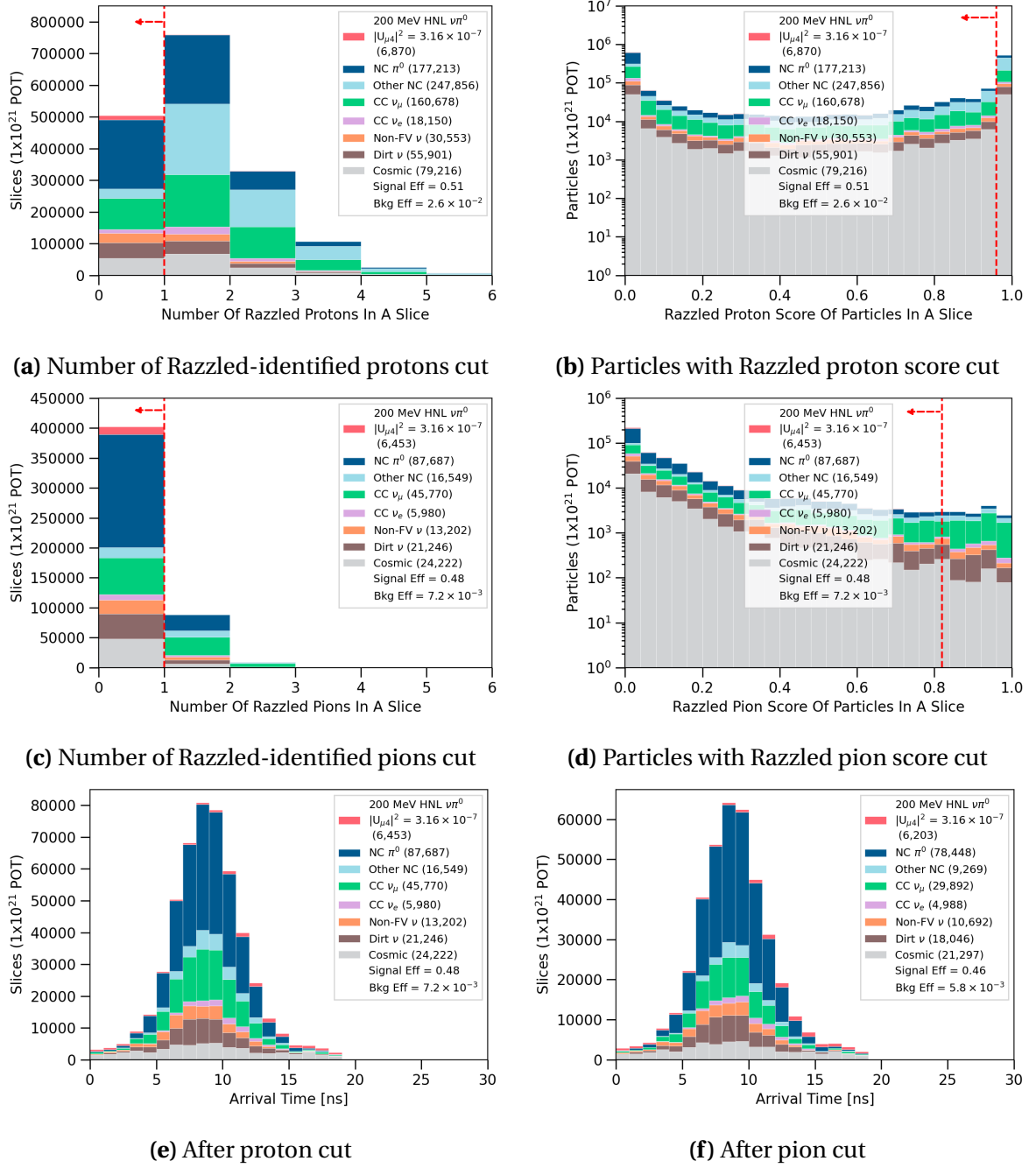
(a) Number of Razzled-identified protons cut

(b) Particles with Razzled proton score cut

(c) Number of Razzled-identified pions cut

(d) Particles with Razzled pion score cut

(e) After proton cut

(f) After pion cut

Fig. 9.7 Proton (top) and pion cuts (middle) and the beam bucket distributions after the cuts (bottom).

conversion gap and the shower dE/dx. The conversion gap describes the gap between the interaction vertex and the start of the shower, where electron showers start immediately at the vertex but photon showers might propagate away from the vertex before showering. The dE/dx describes the charge distribution per unit length, such that the dE/dx of a

photon shower is twice that of an electron shower since a photon shower is made up of a pair of electron-positron showers. Both these shower characteristics are provided during the training of the Razzled BDT for classifying photons and electrons.

The Razzled electron score is examined for the primary shower that deposits the most energy in a slice. The cut is demonstrated in Fig. 9.8, where only slices containing primary showers with a Razzled electron score < 0.96. The rejected slices are clearly-identified CC ν_e showers with high Razzled electron scores. This is a very soft cut compared to the previous track removal cuts since showers from CC ν_e interactions and showers from HNLs are very similar to each other. The cut rejects 31% of the remaining $\sim 5,000$ CC ν_e slices while minimally reduces HNL slices by only 3%.

9.4.2 Track Score Cut

To further reject backgrounds containing showers, careful considerations were taken into developing cuts by separating the shower topology into subsets. As previously stated, di-photon showers from HNLs can result in either a single shower topology or multiple shower topology. Thus, two cases can be considered when applying cuts: (1) slices containing only one shower and (2) slices containing two or more showers. The distribution of signal and background slices in the phase space of the cut variable vary differently between the two cases This results in a different signal-to-background ratio across the distribution for each case. From this cut onwards, individual cut is examined per case to optimise the efficiency of background rejection and signal selection.
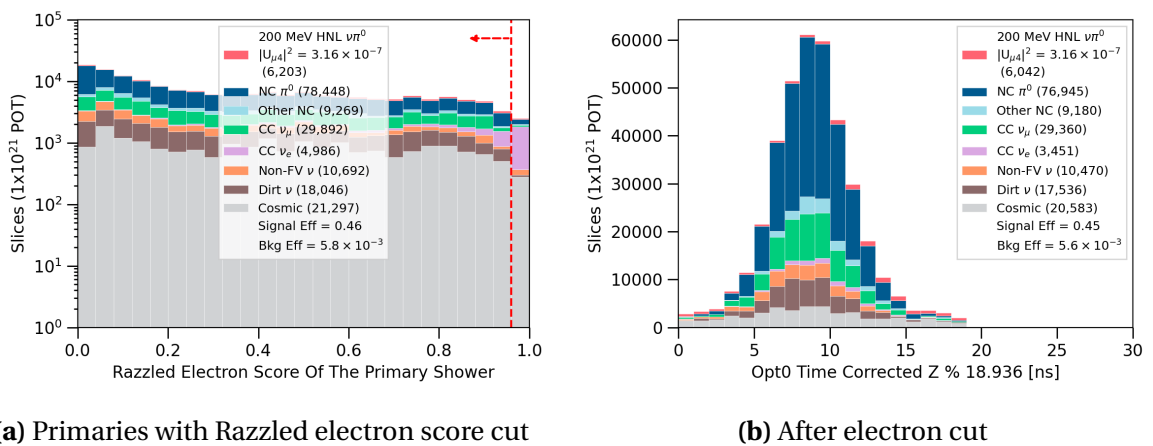

(a) Primaries with Razzled electron score cut   **(b)** After electron cut

Fig. 9.8 Electron cut (top) and the beam bucket distribution after the cut (bottom).

(a) 1 Shower Case: Track score cut

(b) 2+ Showers Case: Track score cut

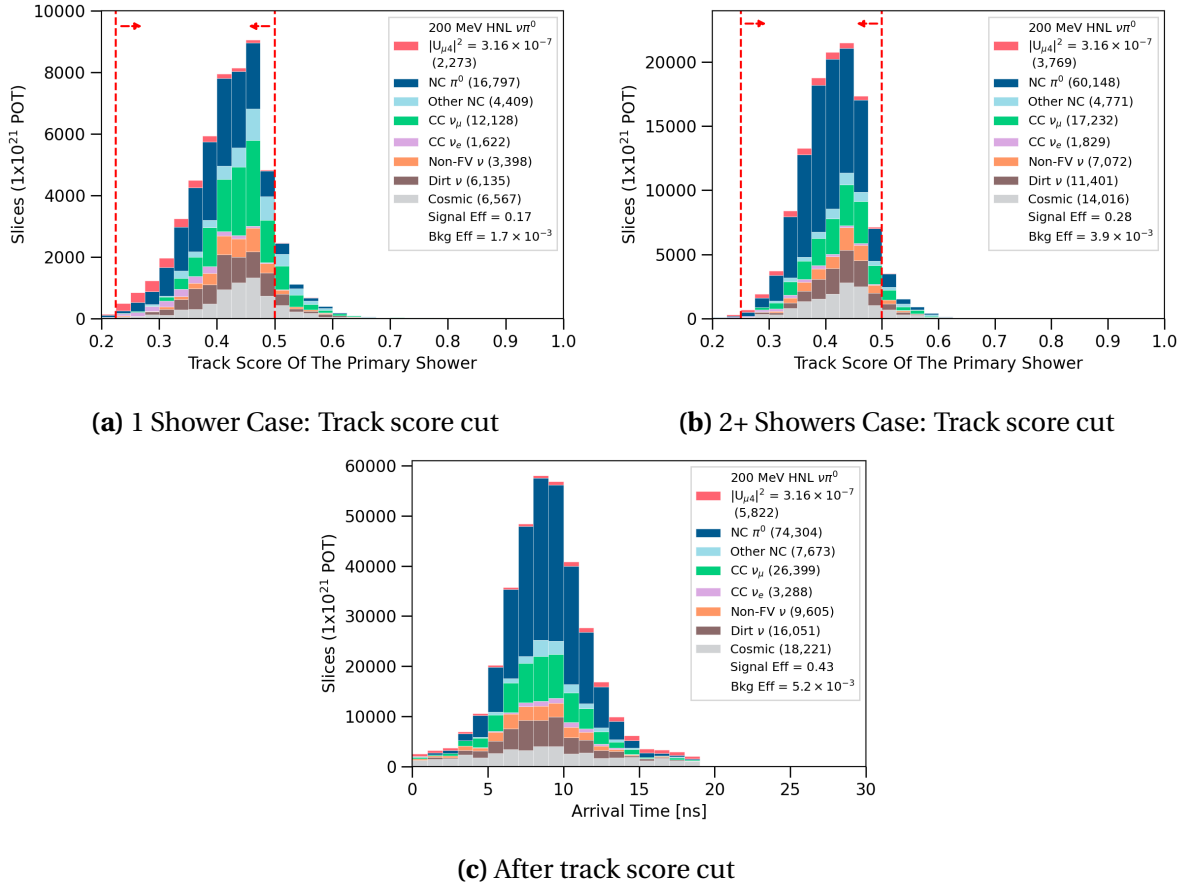**(c)** After track score cut

Fig. 9.9 Track score cut (top) and the beam bucket distribution after the cut (bottom).

The second cut of the HNL shower selection employs the track-shower separation BDT, that outputs a track score to a particle indicating if it is track-like or shower-like (See Section 6.2.4). The track score is examined for the primary shower that deposits the most energy energy in a slice. Fig. 9.9 displays the track score distribution of primary particles for the two cases of slices containing 1 shower and 2+ showers. For both cases, the remaining primary particles are already shower-like since the track score concentrates in the region < 0.5.

The cut sets the upper boundary of the track score at 0.5 for both cases, to reject any primary particles leaning towards track-like. On the other hand, the cut on the lower boundary of track score aims at trimming some shower-like backgrounds. The cut is optimised for each case depending on individual signal-to-background distribution. A more lenient cut is applied for the single shower case selecting the primary shower with a track score of ≥ 0.225. The cut is tightened up for the multiple showers case for better background rejection, requiring the primary shower to have a track score of ≥ 0.25. The resulting beam

bucket distribution is depicted in Fig. 9.9c, showing a reduction of 3-16% across different SM neutrino interaction types. The signal selection efficiency only reduces from 45% to 43%.

9.4.3 Calorimetry Cut

The third cut of the HNL shower selection targets the highly energetic aspect of the showers resulting from HNL decays compared to SM neutrino showers. Outputs from the flash-to-slice matching process are examined, particularly the fraction variable defined in Eq. 6.4 in Section 6.4.1. For a matched flash, the fraction describes the level of agreement between L_Q and L, where L_Q is the number of PhotoElectrons (PEs) predicted from the reconstructed charged and L is the number of PEs measured by PMTs. A large disagreement indicates poor reconstruction, whether under or overestimation in light prediction or non-coincident cosmic backgrounds.

The fraction is useful to identify showers originated from HNLs due to their boosted topology. Very forward-going HNL showers are likely to overlap and clustered as a single shower merged from multiple showers by Pandora. The reconstructed energy from deposited charge of the merged HNL shower tends to be much higher than that for SM neutrinos. The number of PEs predicted from the reconstructed charge L_Q is therefore also likely to be overestimated compared to the number of PEs measured by PMTs L.

The overestimation is demonstrated in Fig. 9.10, where it can be seen that HNL slices mainly concentrate in the region $\frac{(L_Q - L)}{L} \geq 0$. The calorimetry cut exploits this feature and is optimised for the single shower as well as the multiple shower case. For slices containing a single shower, the requirement on the fraction is between -0.1 and 0.4 to select well-predicted showers with the fraction centred around 0, as well as overestimated showers with the fraction > 0. For slices containing multiple showers, the requirement on the fraction is restricted to between 0.04 and 0.3 to strictly select only overestimated showers, rejecting backgrounds more aggressively. The beam bucket distribution after the cut is shown in Fig. 9.10c, demonstrating the effectiveness of the cut as the background efficiency decreases by a whole order of magnitude from $\mathcal{O}(10^{-3})$ to $\mathcal{O}(10^{-4})$. Meanwhile, the signal selection efficiency only decreases from 43% to 35%.

9.4.4 Theta Angle Cut

The fourth cut exploits the topology of the forward-going HNL showers such that their angles with respect to the beam direction, referred to as *theta angles,* are small. Fig. 9.11a shows
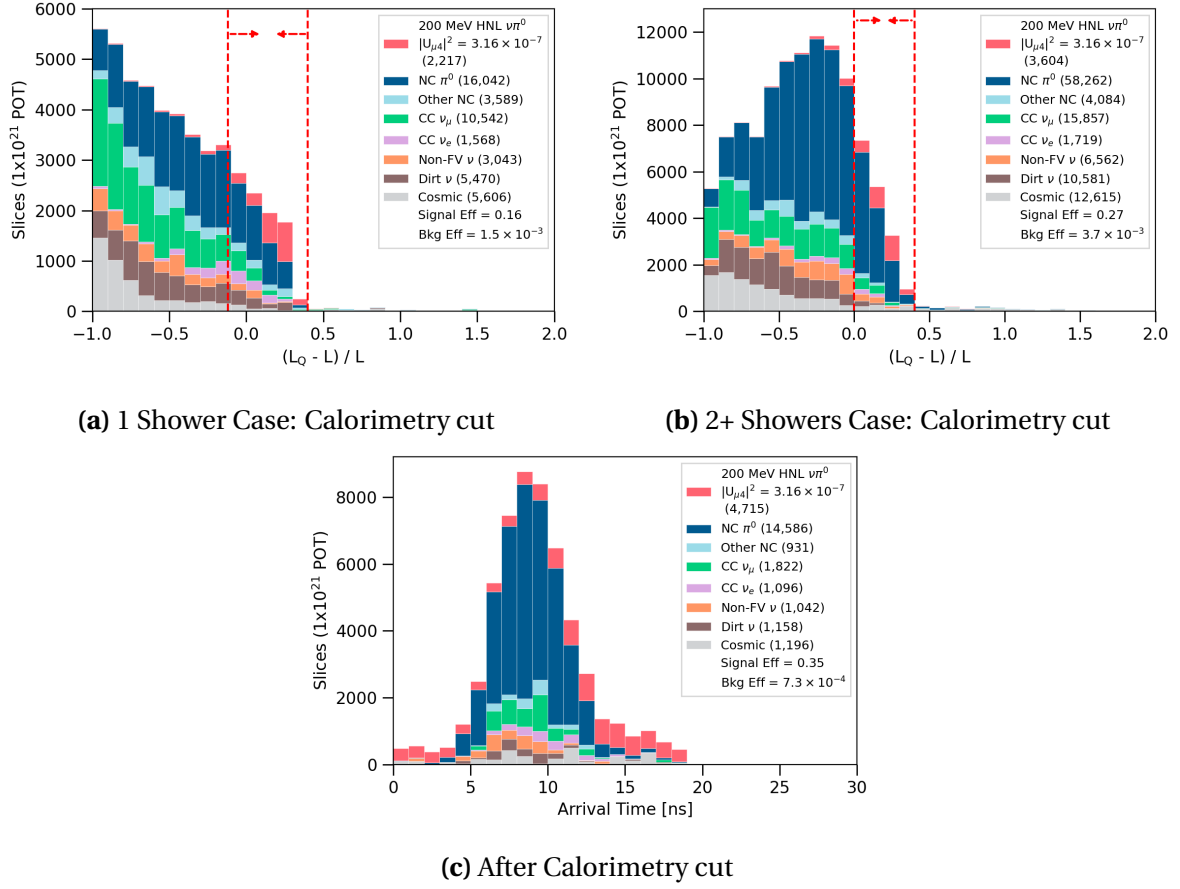
(a) 1 Shower Case: Calorimetry cut

(b) 2+ Showers Case: Calorimetry cut

(c) After Calorimetry cut

Fig. 9.10 Calorimetry cut (top) and the beam bucket distribution after the cut (bottom).

the angular distribution for slices containing a single shower. In this case, the signals are mainly highly energetic and boosted di-photon showers reconstructed as a single merged and beam-collimated shower. As a result, their theta angles concentrate in the region $< 25°$. An aggressive selection of $< 25°$ can be placed without compromising signal efficiency.

Fig. 9.11b shows the theta angle distribution for slices containing multiple showers. In this case, HNL showers are less boosted and more likely to result in separated showers. Their theta angles with respect to the beam are larger compared to the single shower case. To preserve signal selection efficiency, a widened selection of $< 30°$ is applied.

Fig. 9.11c shows the beam bucket distribution after applying the cut. The theta angle cut effectively rejects any shower-like backgrounds that are not beam-collimated, resulting in a reduction across all SM neutrino interaction types. This is a very impactful cut given that the background efficiency decreases by half from 7.3×10^{-4} to 3.6×10^{-4}. The signal efficiency of HNL slices only drops by 2%.

9.4.5 Invariant Mass Cut

The final cut of the HNL shower selection exploits the fact that di-photon showers originate from a π^0 decay, allowing for the reconstruction of π^0 invariant mass, m_{π^0}. For slices containing multiple showers, the invariant mass can be reconstructed using the reconstructed momenta of any two showers combination in the slice. For two massless photon showers with an opening angle α and a total energy E_1 and E_2 respectively, m_{π^0} is computed as:

$$m_{\pi^0} = \sqrt{2 E_1 E_2 \times (1 - \cos\alpha)}. \tag{9.4}$$

For a given slice, the π^0 invariant mass was reconstructed for all combinations of two showers, and the best mass was considered for the cut. The cut is illustrated in Fig. 9.12, where the solid red line indicates the π^0 mass of 135 MeV. A cut is applied to select slices corresponding to a reconstructed invariant mass of 300 MeV or less. This rejects any slices
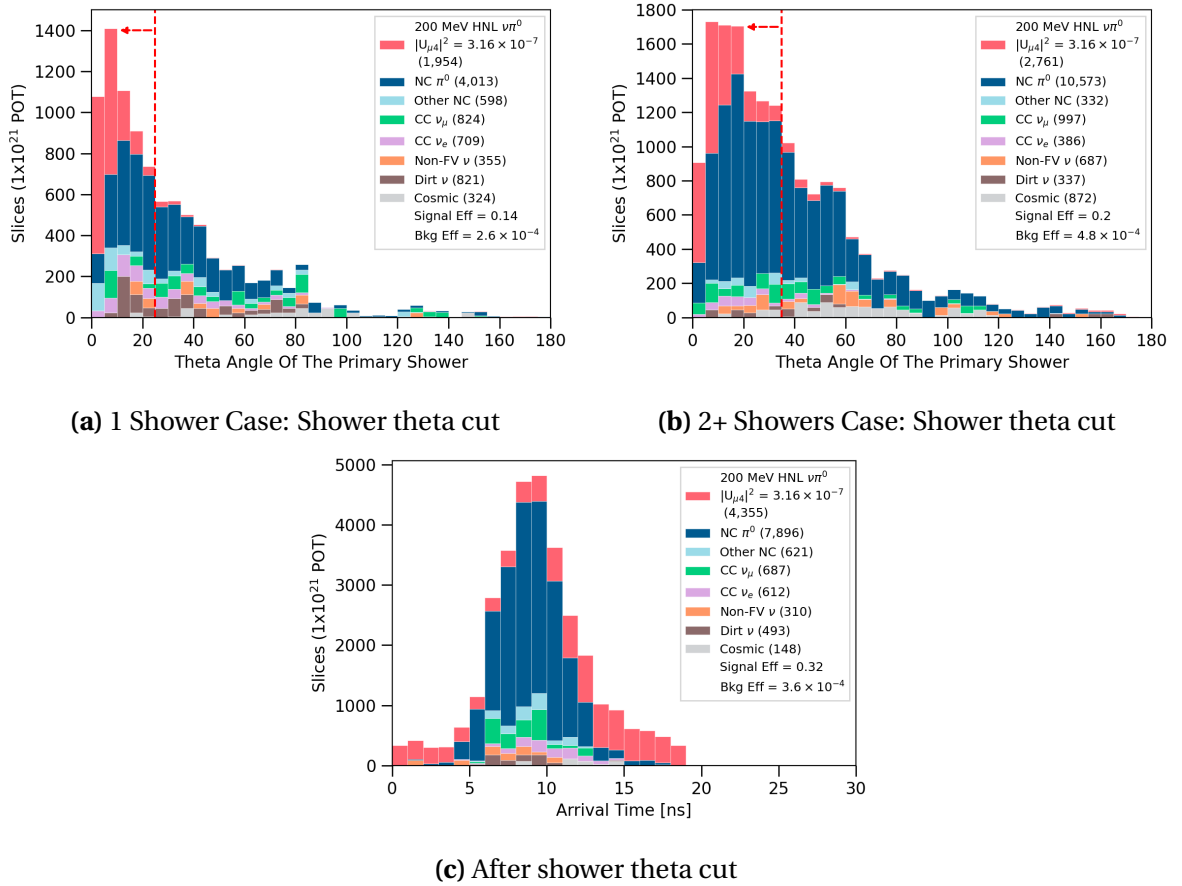

(a) 1 Shower Case: Shower theta cut

(b) 2+ Showers Case: Shower theta cut

(c) After shower theta cut

Fig. 9.11 Shower theta cut (top) and the beam bucket distribution after the cut (bottom).

with a poorly reconstructed π^0 mass, which could be due to backgrounds from SM neutrino interactions such as CC ν_μ, Other NC, Non-FV and dirt as well as energetic cosmic rays. However, poor shower reconstruction can also result in di-photon showers from π^0 getting mistakenly rejected by this cut, as it is evident that some NC π^0 interactions and HNL signals are affected. This cut reduces both signal and background slices by than 3%.
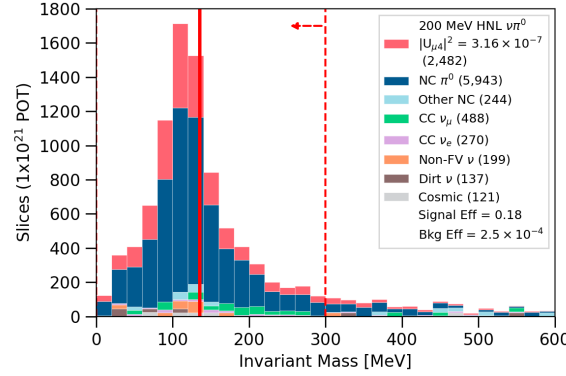

Fig. 9.12 Invariant mass cut applied to the multiple showers case.

9.5   Final Selection Result

Fig. 9.13a shows the beam bucket distribution after the whole selection procedure. The background efficiency shown in the plot is 3.3×10^{-4}, demonstrating the extreme background rejection achieved for this analysis amounting to four orders of magnitude. Meanwhile, the signal selection efficiency is still well-preserved as 30%.

The peak region of the bucket is dominated by the primary background from NC π^0 interactions which have proven to be a very tricky background to remove due to their similarity with HNL showers. A combination of CC ν_μ and Other NC remain even though they were not considered to be a background at the beginning. These interactions likely undergo deep inelastic scattering, producing shower-like products like π^0 or $e^\pm$. Moreover, a fraction of CC ν_e interactions persists as they can produce a single shower topology. In addition, some backgrounds from Non-FV and dirt neutrino interactions can still be seen, since their products can propagate to the FV and deposit energy. Finally, cosmic muons are almost fully rejected is due to aggressive cuts.

The multi-binned analysis for limits setting depends on the signal-to-background ratio per bin, of which signal-rich bins drive the limits. Fig. 9.13c zooms into the first and last 4 bins of the beam bucket distribution. These are of the highest purity of the entire

histogram and contribute towards the final sensitivity significantly more than bins located at the peak region. A *timing cut* might be applied to select only these bins, which would result in a background rejection efficiency increasing from $\mathcal{O}(10^{-4})$ to $\mathcal{O}(10^{-6})$ while still maintain a signal efficiency of 10%. However, the cut is not formally applied as part of the selection procedure, but to highlight the importance of these edge bins due to their excellent signal-to-background ratio. The timing cut is discussed further in Chapter 10.

To better understand the sensitivity dependency on the signal-to-background ratio, two selection procedures were developed. The selection demonstrated up until this point is referred to as *the lenient cut*. An additional more aggressive cut, referred to as *the stringent cut,* was developed by tightening the two most impactful cuts on calorimetry and theta angle. The resulting beam bucket distribution for the stringent cut is plotted in Fig. 9.13b and 9.13d for the entire distribution and only the edge bins respectively. The key difference between these two cuts is that the lenient cut retains more signals however at a lower purity. Meanwhile, the stringent cut results in higher purity at the cost of signal efficiency. The two selections are summarised in Table 9.1.

Fig. 9.14 shows the signal selection and background rejection efficiency cut by cut. The signal selection efficiency is plotted using the left axis in pink and the background rejection efficiency is plotted using the right axis in blue. It is important to note that the right axis is in the logarithm scale as the background rejection is very aggressive. The band of signal selection efficiency corresponds to the efficiency across the entire mass range of HNLs from 140 to 260 MeV, with efficiency increasing with mass. The selection differs from the calorimetry cut onwards, where the lenient cut is shown in pink and the stringent cut is shown in red. Overall, the most significant cuts are the muon/proton/pion cut for track removal, followed by the calorimetry and theta angle cut by exploiting the boosted topology of HNLs that significantly reject backgrounds without compromising signal efficiency.

9.6   Study of Timing Resolution Improvement

The timing cut described above demonstrates the importance of a high precision timing reconstruction in this analysis and thus, a study was carried out to understand the smearing contributors to the beam bucket distribution. Fig. 9.15 illustrates several factors that can smear the arrival time of a SM neutrino at SBND, and consequently smearing the Gaussian shape of the beam bucket. The intrinsic Gaussian width of the proton bucket from the Booster synchrotron is 1.308 ns, as shown by the brown arrow. This structure is then smeared out due to the Time of Flight (ToF) of secondary mesons, as shown by the blue
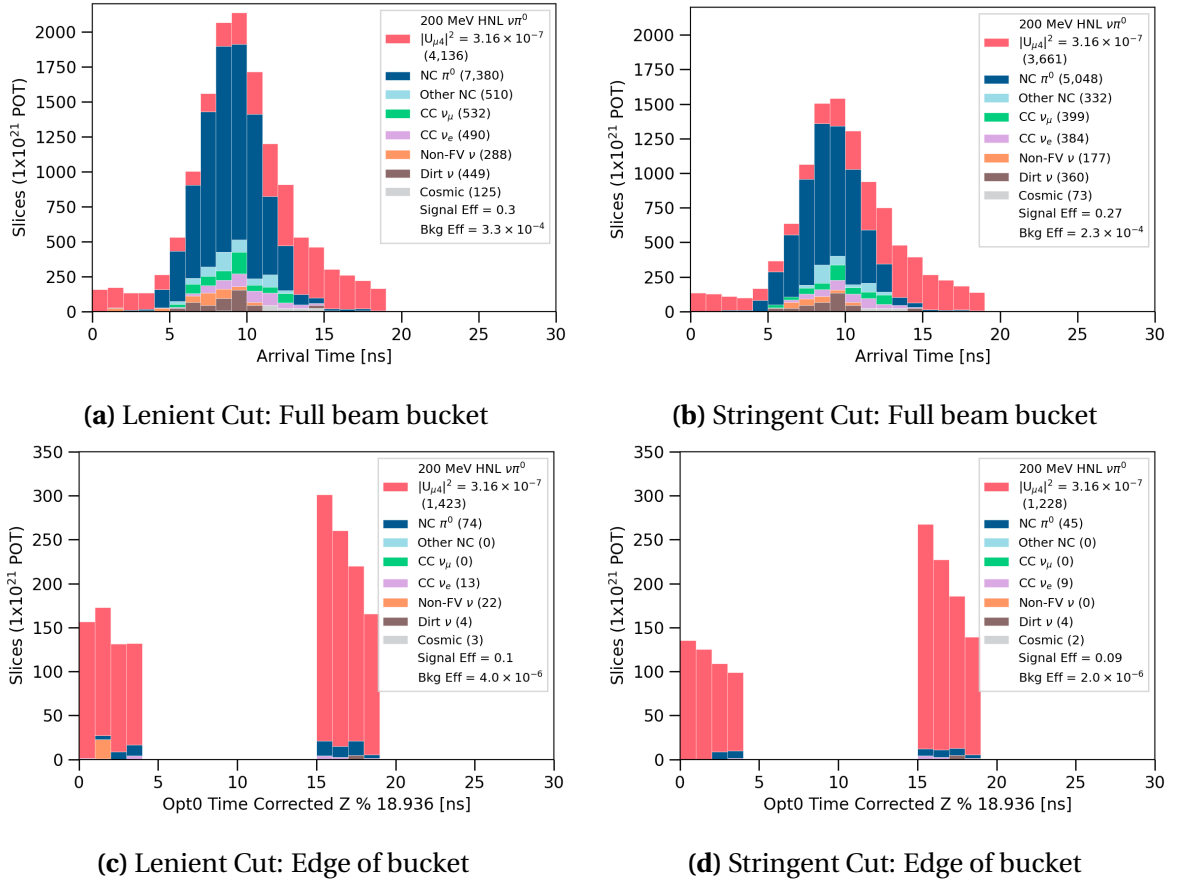
(a) Lenient Cut: Full beam bucket

(b) Stringent Cut: Full beam bucket

(c) Lenient Cut: Edge of bucket

(d) Stringent Cut: Edge of bucket

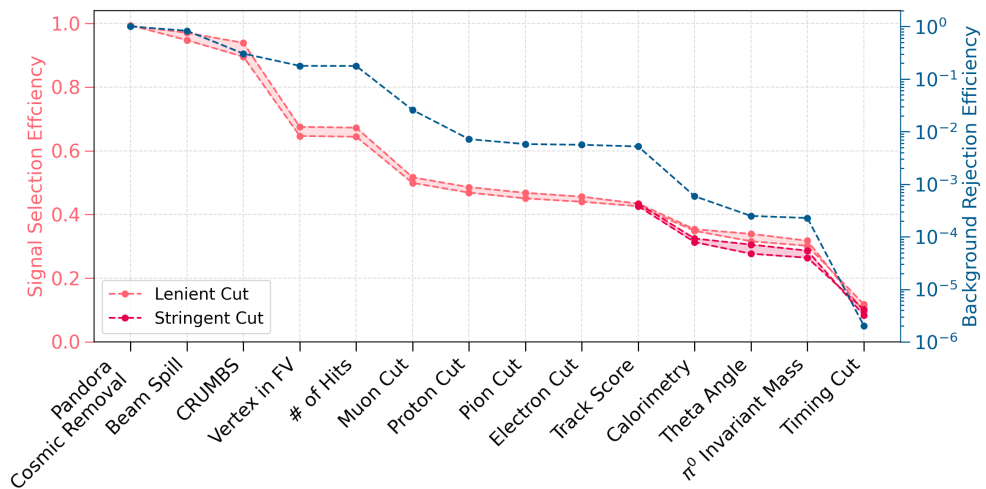**Fig. 9.13** The beam bucket distributions after the lenient (left) and stringent cut (right).

Fig. 9.14 Summary of signal (left axis) and background efficiencies (right axis).

Table 9.1 Table summarising the lenient and stringent selection.

	Common Cut	
Cosmic Removal:		
Slice reconstructed by Pandora as a neutrino	True	
Flash time inside the beam spill	[0.367, 1.967] μs	
CRUMBS score	≥ 0	
SM Neutrino Removal:		
Reconstructed vertex inside the FV	True	
# of hits in the primary shower	≥ 50	
# of Razzled muons	0	
Razzled muon score of particles in a slice	< 0.04	
# of Razzled protons with KE > 32.7 MeV	0	
Razzled proton score of particles in a slice	< 0.96	
# of Razzled pions with KE > 31.2 MeV	0	
Razzled pion score of particles in a slice	< 0.82	
HNL Shower Selection:		
Razzled electron score of the primary shower	< 0.96	
Track score of the primary shower		
1 shower case	$0.225 <$ score < 0.5	
2+ shower case	$0.250 <$ score < 0.5	
	Lenient Cut	**Stringent Cut**
(L_Q - L) / L fraction of a slice		
1 shower case	$-0.12 <$ frac < 0.40	$-0.10 <$ frac < 0.40
2+ showers case	$0.00 <$ frac < 0.40	$0.04 <$ frac < 0.30
Theta angle of the primary shower		
1 shower case	$\leq 25°$	$\leq 20°$
2+ showers case	$\leq 35°$	$\leq 30°$
Invariant mass of any 2 showers in a slice	≤ 300 MeV	
Timing Cut *(applied when setting limits):		
Arrival time within the beam bucket	$0 \leq t \leq 4$ and $15 \leq t \leq 19$	

arrow. Moreover, the ToF of the tertiary SM neutrinos from the production location to the detector further smears the Gaussian, as shown by the pink arrow. Once the neutrino arrives at the detector, two additional smearing factors need to be considered. The first one is its ToF inside the detector, as shown by the purple arrow. The second one is the ToF of the photon from the production to the detection location, assuming the photon production location is close to the interaction vertex, as shown by the green arrow.
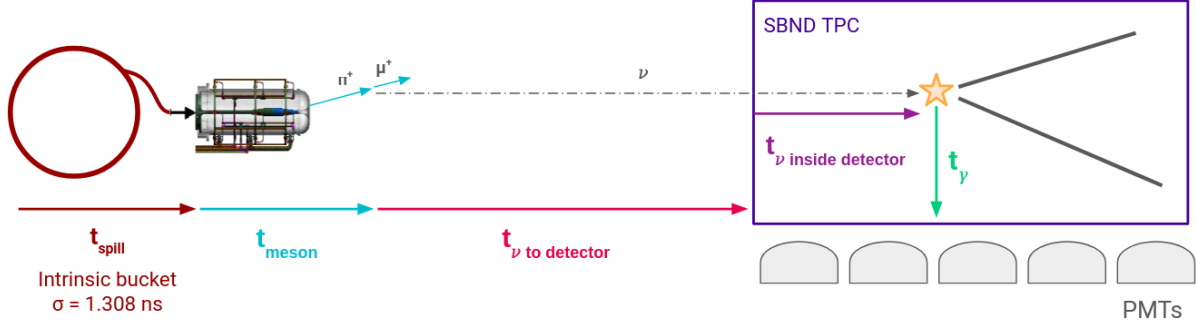

Fig. 9.15 Diagram showing different smearing contributors to the beam bucket distribution.

A beam bucket distribution was plotted using *true* variables to better understand the impacts of the smearing factors, where true indicates that no detector simulation and reconstruction were applied. The truth beam bucket of SM neutrinos that arrive at the front face of SBND is shown in the left of Fig. 9.16. This distribution was computed using the true interaction time at the vertex, marked by the yellow star in Fig. 9.15, and then corrected for the SM neutrino ToF inside the detector, depicted by the purple arrow. In the truth phase space, it is evident that the combination of the ToF of mesons and of SM neutrinos to the detector smears the Gaussian width by a negligible amount from 1.308 ns to 1.37 ns.

The reconstructed beam bucket distribution of SM neutrinos is plotted in the right of Fig. 9.16. As detailed in Section 9.1.4, the beam bucket is reconstructed using the flash time that is matched to a slice (See Sections 6.3.1 and 6.4.1). The flash time was reconstructed using the prompt light occurring in the first 30 ns window such that the scintillation location is close to the interaction vertex. The flash time was also corrected for the photon propagation time from the production location to PMTs, depicted by the green arrow in Fig. 9.15. Then, the SM neutrino ToF inside the detector, as shown by the purple arrow, was corrected by applying a shift from the reconstructed vertex z-position to $z = 0$ at the detector's front face. As a result, the reconstruction depends on 3 variables: (1) the matching of flash-to-slice, (2) the flash time and (3) the slice vertex. Each of these variables has its own reconstruction uncertainty, thus, the reconstructed beam bucket is even more smeared.

Comparing the two beam bucket distributions in Fig. 9.16, the Gaussian mean is shifted by 1.82 ns from 7.44 to 9.26 ns. This includes a shift of 1.45 ns introduced by the light reconstruction [2]. The rest might be due to the slice vertex reconstruction and/or the slice-to-flash matching. Additionally, the Gaussian width is smeared from 1.37 ns to 2.26 ns. This width smearing is detrimental to the HNL search since it results in more SM neutrinos at the edge of the beam bucket, reducing the signal-to-background ratio in this region.
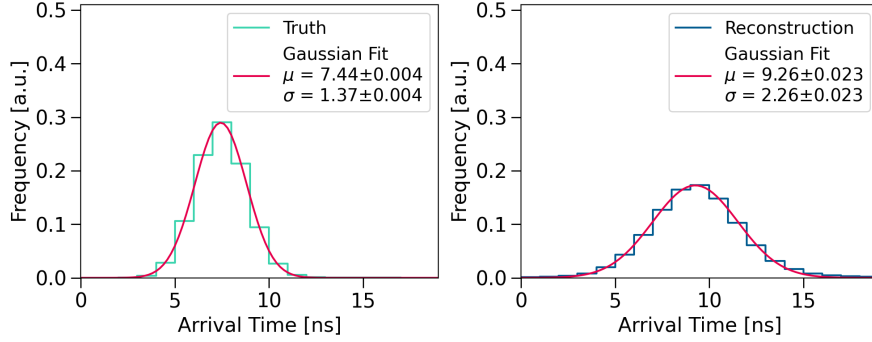

Fig. 9.16 The beam bucket distribution of SM neutrinos in the truth phase space (left) and in the reconstruction phase space (right).

This motivates the assessment of sensitivity assuming a better timing reconstruction. The two assumptions of the beam bucket distribution reconstructed with an improved timing resolution are as follows:

1. A shifted Gaussian mean of 1.45 ns,

2. A smeared Gaussian width of 1.73 ns.

The first assumption is motivated by the impact of the light reconstruction in SBND reported by Ref. [2]. The second assumption is motivated by the MicroBooNE experiment reporting on their intrinsic timing resolution [1]. Although ambitious, it is an achievable goal for SBND to have a reconstructed timing resolution < 2 ns, given that SBND employs a similar detector technology to MircoBooNE. Moreover, Chapter 7 details the excellent timing performance of the SBND data acquisition and the preparation that already took place to achieve better timing resolution. Particularly, the SPEC-TDC device, as discussed in Section 7.1.2, records important timing information on the trigger and beam arrival that can only improve downstream reconstruction once incorporated.

For modelling the background using true variables under these assumptions, only SM neutrinos are considered and not cosmics for simplicity. The truth beam bucket distribution of SM neutrinos was smeared with the two assumptions, referred to as *smeared truth*. Fig.

9.17 shows the truth, smeared truth and the reconstructed distribution after selection, all normalised to the same area for direct comparison. The left figure shows the truth distribution without any smearing applied with a width of 1.37 ns. The middle figure shows the smeared truth distribution with the assumed width of 1.73 ns. The right figure shows the reconstructed distribution after applying the lenient selection with a width of 1.99 ns. It is important to note the difference between the reconstructed beam bucket before and after selection such that the distribution after selection has a less shifted Gaussian mean and a smaller Gaussian width. This is due to cuts having non-uniform effects on the distribution. It was observed in this work that different topologies, for instance tracks and showers, have different timing reconstruction resolutions.
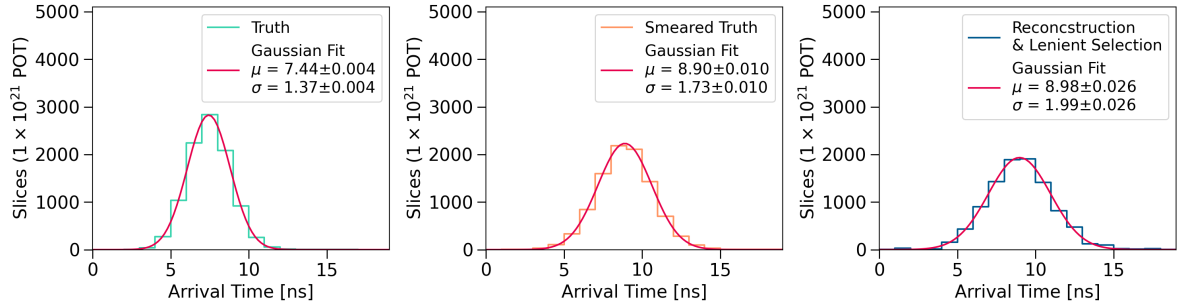

Fig. 9.17 The beam bucket distribution of SM neutrinos in truth (left), smeared truth (middle) and reconstruction after selection (right).

For modelling the signal, the same smearing assumptions are applied to the truth timing distribution of HNLs.

Unlike the background modelling approach of normalising the same area, an efficiency of 30% was applied to the HNL truth beam bucket distribution to account for the combined effects of reconstruction and selection. Fig. 9.18 shows the beam bucket distribution of SM neutrinos and HNLs across the phase space of truth, smeared truth and reconstruction after selection for comparison. The smeared truth distribution shows a higher signal-to-background ratio particularly for the bins at the edge of the bucket compared to the reconstructed distribution. This smeared truth distribution is also used for determining the sensitivity alongside the reconstructed distributions, as this is to assess the impacts of timing resolution improvement.
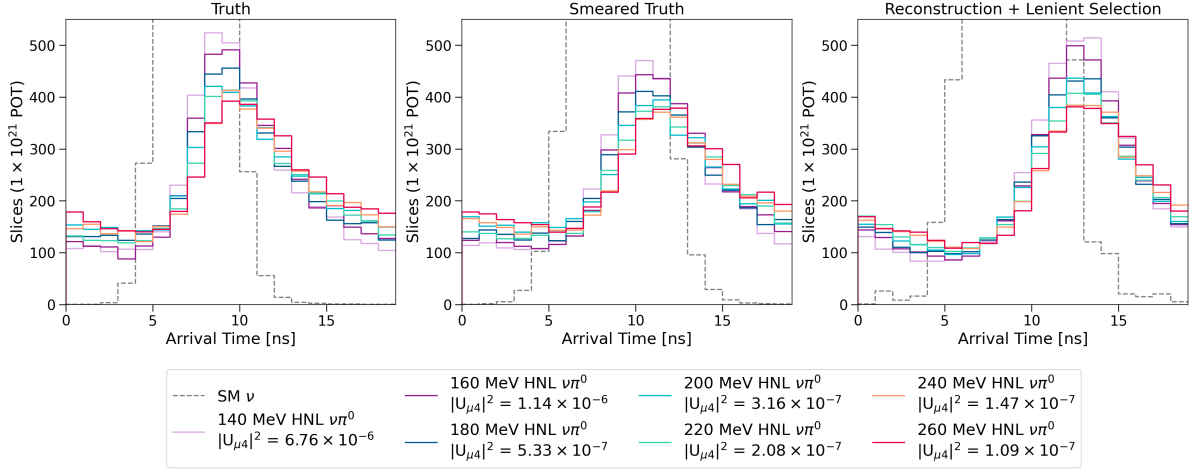
Fig. 9.18 The beam bucket distribution of SM neutrinos and HNLs in truth (left), smeared truth (middle) and reconstruction after selection (right).

9.7 Concluding Remarks

The selection of HNLs using MC samples is provided in this chapter, as a procedure to identify HNL signals from SM neutrino and cosmic backgrounds. Two selection procedures on reconstructed variables are presented, the lenient and stringent cut, with the stringent rejecting backgrounds more aggressively than the lenient. Both exploit the highly energetic and forward-going features of HNL showers to achieve an excellent background rejection without compromising signal efficiency.

The resulting background efficiency is in the order of $\mathcal{O}(10^{-4})$ while the signal selection efficiency still maintains at 30%. When considering only bins at the edge of the beam bucket distribution, or the so-called *timing cut*, the background efficiency decreases significantly by two orders of magnitude to $\mathcal{O}(10^{-6})$. Meanwhile, the signal efficiency only decreases from 30% to 10%. This demonstrates that these edge bins contain an exceptional signal-to-background ratio, which is the main factor driving the sensitivity.

Furthermore, a study was motivated to explore the impact on sensitivity if the timing reconstruction is achieved. The study resulted in a beam bucket distribution acquired by smearing true variables, assuming it is reconstructed with an improved timing resolution. All three beam bucket distributions, from both the lenient and stringent cut on reconstructed variables and from the smeared truth variables, are used for studying the sensitivity to HNLs in Chapter 10 next.

147

References

[1] P. Abratenko et al. (MicroBooNE Collaboration), "First demonstration of $\mathcal{O}(1 \text{ ns})$ timing resolution in the MicroBooNE liquid argon time projection chamber", Phys. Rev. D **108**, 052010 (2023).

[2] P. Abratenko et al. (SBND), "Scintillation Light in SBND: Simulation, Reconstruction, and Expected Performance of the Photon Detection System", (2024).