# Signal Region Optimization for the search of Right Handed Neutrinos

Srishti Patil

## 1 Introduction

The Standard Model of particle physics (SM) along with the theory of General Relativity (GR) can describe almost everything observed in nature, barring a few experimental and observational facts like baryon assymetry in the universe and the composition and origin of Dark Matter. *Right handed neutrinos* (RHN), if they exist, could be responsible for a lot of these unexplained phenomena. They are promising Dark Matter candidates as well.

This project is a part of the search for RHN using data samples collected by the CMS experiment at the LHC in 2016 [Table 1], corresponding to an integrated luminosity of 36 $\text{fb}^{-1}$. We employ a technique called multilepton analysis [Section 2] for the search and the project focuses on a part of the process called Signal Region (SR) optimization. The results obtained are used to produce limits on the production of right handed neutrinos.

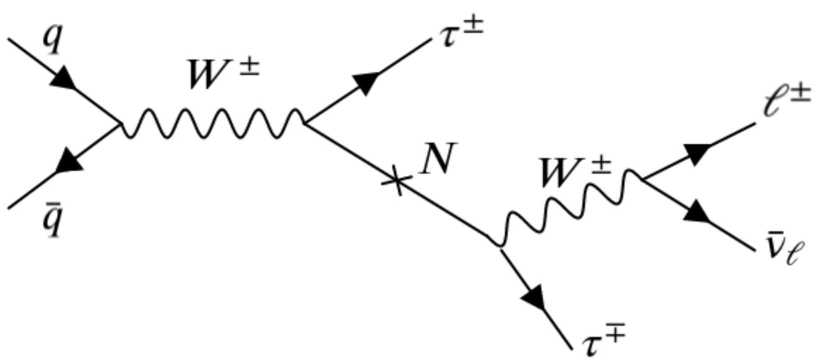



*Figure 1:* Illustrative leading order Feynman diagram for the production of RHN and its subsequent decay that may result in a multilepton final state. The taus can decay either hadronically or leptonically, giving us different channels to analyze that have different background profiles

## 2 Overview

The final state of an event (collision) is the set of particles seen by the detector. A lot of processes Beyond the Standard Model (BSM) produce final states with multiple leptons. Multilepton

| Sample (RHN Mass Point) | Total Number of Events (n) | Cross-Section (xsec) (pb) | Luminosity ($fb^{-1}$) |
|---|---|---|---|
| 100 GeV | 299,200 | 0.830 | 360.482 |
| 150 GeV | 286,200 | 0.119 | 2405.042 |
| 400 GeV | 200,000 | 0.00272 | 73529.411 |

*Table 1:* Data samples used for this analysis. Three mass points are considered and their respective sample luminosities are calculated using the formula $L_{sample}$ = $n_{sample}$ / $xsec_{sample}$

Analysis is a useful technique for studying such processes. Since a lot of SM processes also produce multilepton final states (called background), we use this strategy to distinguish signal from background.

The model considered here [Figure 1] includes a right handed neutrino ($N$) that decays into a $W$ boson and a $\tau$, which decay further, giving rise to multilepton final states. For this analysis, we refer to electrons and muons collectively as light leptons ($\ell$). Events are then categorized by the multiplicity of light leptons and taus ($\tau$) in these final states, into the following channels: $3\ell$, $2\ell 1\tau$, $1\ell 2\tau$ and $3\tau$.

To begin with, we generate the RHN acceptance cutflow table [Section 6, Table 6] to understand the process and how various triggers and selections affect the acceptance. Next, we attempt to understand the process of RHN production and decay at the gen-level ["'ref section"'] (*i.e.* from simulations). Guided by the conclucions drawn from the study of these gen-level plots, we try to carve out a Signal Region [Section 5] from the phase space which will give us good signal sensitivity. We use signal significance as a metric to compare signal regions. We do this for two channels: $2\ell 1\tau$ and $3\ell$.

# 3 Object and Event Selections

To better identify particles significant to our analysis, we apply certain basic preselections and triggers to the data. Object selections are listed in Table 2 and Table 3. Triggers are listed below. A few corrections like MET filters (on data and MC both), tau energy scale corrections and tau reco scale factors are also applied.

Triggers:

- Leading muon $p_T > 26$ GeV
- Leading electron $p_T > 30$ GeV
- For Single Muon PD, HLT_IsoMu24 & HLT_IsoTkMu24 (in data only)
- For Single Electron PD, HLT_Ele27_WPTight_Gsf (in data only)

| Muons | Electrons |
|---|---|
| • $p_{\mathrm{T}} > 10, \lvert\eta\rvert < 2.4$<br>• Prompt ($d_{xy} < 0.05, d_z < 0.1$ cuts)<br>• POG Medium ID<br>• Relative dB-corrected isolation tight WP ($< 0.15$ in $\Delta R = 0.4$) | • $p_{\mathrm{T}} > 10, \lvert\eta\rvert < 2.5$<br>• Prompt ($d_{xy} < 0.05(0.1), d_z < 0.1(0.2)$ in barrel(endcap) region)<br>• Cut-based Medium ID<br>• Relative rho-corrected isolation Medium WP (included in ID) |

*Table 2:* Object selections for light leptons

| Taus | |
|---|---|
| MVA ID (old) | Deep ID (new) |
| • $p_{\mathrm{T}} > 20, \lvert\eta\rvert < 2.3$<br>• Prompt ($d_z < 0.2$ cut)<br>• Old decayModeFinding, 1-prong & 3-prong<br>• '2017v2' ID<br>• againstElectron & againstMuon discriminators, loose WP<br>• Cleaning against tight ID light leptons ($\Delta R > 0.4$) | • $p_{\mathrm{T}} > 20, \lvert\eta\rvert < 2.3$<br>• Prompt ($d_z < 0.2$ cut)<br>• New decayModeFinding, 1-prong & 3-prong<br>• Tau_idDeepTau2017v2p1VSe $\geq 15$ & Tau_idDeepTau2017v2p1VSmu $\geq 3$ (Loose WP)<br>• Tau_idDeepTau2017v2p1VSjet > 31 (Tight WP) |

*Table 3:* Object selections for taus. A comparision between the two object IDs is given in Section 6.

The events used for analysis are selected only if they pass certain conditions (selections) listed in Table 4. An event is selected for a channel only if it fails the selections for all the previous channels. The order of priority (event selection flow) is illustrated in Figure 2.

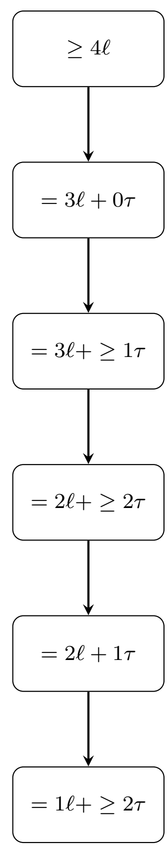
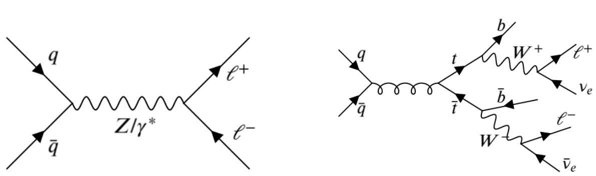



*Figure 2:* Event selection flow

# 4 Background

Many SM processes give rise to multilepton final states, and our objective is to distinguish these processes (background) from the signal (RHN decay). The method used for background estimation in this analysis is Monte-Carlo based. All the background processes considered are listed in Table 5. $WZ$ and $ZZ$ are prompt backgrounds while DY and $t\bar{t}$ are fake (*i.e.* one of the leptons is faked by another particle/jet).




*(a)* DY *(b)* $t\bar{t}$

*Figure 3:* Leading order Feynman diagrams for fake background processes. *(a)* The Drell-Yan process. The third lepton can be faked by another particle/jet. *(b)* One of the possible decay products of the $t\bar{t}$ process

| Event Selections |
|---|
| $3\ell$ |
| • Leading light lepton ($e/\mu$) passing triggers<br>• Altleast two more light leptons ($e/\mu$), $p_{\mathrm{T}} > 10, 10$ GeV<br>• $\Delta R > 0.4$ among the three leptons<br>• Minimum invariant mass of same flavor (SF) light lepton pair $> 12$ GeV<br>• Minimum $\Delta R$ between all leptons in the event $> 0.4$<br>• Triggers (in data only) |
| $2\ell 1\tau$ |
| • Event should fail $3\ell$ selection<br>• Leading light lepton ($e/\mu$) passing triggers<br>• Another light lepton ($e/\mu$), $p_{\mathrm{T}} > 10GeV$<br>• Minimum invariant mass of same flavor (SF) light lepton pair $> 12$ GeV<br>• Atleast one hadronic tau, $p_{\mathrm{T}} > 20$ GeV<br>• $\Delta R > 0.4$ among the three leptons<br>• Triggers (in data only) |
| $1\ell 2\tau$ |
| • Exactly one light lepton ($e/\mu$), for triggering<br>• Atleast two hadronic taus, $p_{\mathrm{T}} > 20, 20$ GeV<br>- Prompt taus passe tight ID, fake taus pass loose ID<br>• $\Delta R > 0.4$ among the three leptons<br>• Triggers (in data only)<br>• Event should fail all other event ($3\ell, 2\ell + \geq 1\tau..$) selections |

*Table 4:* Event selections for $3\ell$, $2\ell 1\tau$ and $1\ell 2\tau$ channels

| Background | | Total Number of Events (n) | Cross-Section (xsec) (pb) | Luminosity ($\mathrm{fb}^{-1}$) |
|---|---|---|---|---|
| $2\ell 1\tau$ | DY | 89,832,690 | 5765 | 15.582 |
| | $WZ$ | 6,610,401 | 5.052 | 1308.472 |
| | $ZZ$ | 7,547,891 | 1.325 | 5696.521 |
| | $t\bar{t}$ | 24,265,024 | 88.29 | 274.833 |
| $3\ell$ | DY | 89,832,690 | 5765 | 15.582 |
| | $WZ$ | 1,295,229 | 5.052 | 256.378 |
| | $ZZ$ | 1,041,601 | 1.325 | 786.114 |
| | $t\bar{t}$ | 24,265,024 | 88.29 | 274.833 |

*Table 5:* A few background processes and their luminosities. Luminosity is calculated by the formula $L$ = n / xsec

# 5 Signal Region Optimization

The objective of Signal Region optimization is to find a region in the phase space in which the signal is easily distinguished from background. A metric for the same is signal significance, defined as-

$$Significance = \frac{Signal}{\sqrt{Background}}$$

The higher the significance, the better is the signal region. We perform signal region optimization for the $2\ell 1\tau$ and $3\ell$ channels separately because the background profiles for these channels are different.

## 5.1 Understanding the Process with Simulations

To undrestand the general topology of the events, like the $\Delta R$ between leptons and the origin (mother particle) of the particles involved, we plot these quantities using data from simulations.

## 5.2 $2\ell 1\tau$

## 5.3 $3\ell$

# 6 RHN Acceptance Cut-flow

# 7 Results

| | MVA | | | | | | Deep | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Selection | 150GeV | 100GeV | DY | $WZ$ | $t\bar{t}$ | $\mathrm{W+jets}$ | 150GeV | 100GeV | DY | $WZ$ | $t\bar{t}$ | $\mathrm{W+jets}$ |
| Total events ran | 285816 | 298880 | 23237 | 52560 | 24235935 | 57358383 | 285816 | 298880 | 23237 | 52560 | 24235935 | 57358383 |
| Total events | 286200 | 299200 | 23237 | 52560 | 24265024 | 57402435 | 286200 | 299200 | 23237 | 52560 | 24265024 | 57402435 |
| $4\ell$ events | 2 | 4 | 2 | 9 | 855 | 0 | 2 | 4 | 2 | 9 | 855 | 0 |
| $N_{1\ell}$ events | 166649 | 140098 | 23237 | 52560 | 18340314 | 18917151 | 166649 | 140098 | 23237 | 52560 | 18340314 | 18917151 |
| $N_{1\ell\_trigg}$ events | 99912 | 72436 | 21406 | 47589 | 13065106 | 10297682 | 99912 | 72436 | 21406 | 47589 | 13065106 | 10297682 |
| $N_{1\ell\_trigg\_1\ell}$ events | 22994 | 10915 | 21033 | 38399 | 2835198 | 1538 | 22994 | 10915 | 21033 | 38399 | 2835198 | 1538 |
| $N_{1\ell\_trigg\_1\ell p_{\mathrm{T}}10}$ events | 21495 | 9111 | 20680 | 36568 | 2726718 | 396 | 21495 | 9111 | 20680 | 36568 | 2726718 | 396 |
| $N_{1\ell\_trigg\_1\ell p_{\mathrm{T}}10\_1\ell}$ events | 2681 | 857 | 54 | 825 | 16869 | 2 | 2681 | 857 | 54 | 825 | 16869 | 2 |
| $N_{1\ell\_trigg\_2\ell p_{\mathrm{T}}10}$ events ($3\ell$ events) | 2251 | 619 | 37 | 746 | 10002 | 1 | 2251 | 619 | 37 | 746 | 10002 | 1 |
| $N_{1\ell\_trigg\_1\ell p_{\mathrm{T}}10\_1\tau}$ events | 4524 | 1022 | 16333 | 29722 | 1819 | 0 | 3389 | 663 | 11742 | 29813 | 1294 | 0 |
| $N_{1\ell\_trigg\_1\ell p_{\mathrm{T}}10\_1\tau p_{\mathrm{T}}20}$ events ($2\ell 1\tau$ events) | 4280 | 908 | 16329 | 29706 | 1696 | 0 | 3389 | 663 | 11742 | 29813 | 1294 | 0 |
| $N_{1\ell\_trigg\_1\tau}$ events | 23319 | 10796 | 189 | 7478 | 876646 | 7 | 17542 | 7315 | 284 | 7562 | 590115 | 6 |
| $N_{1\ell\_trigg\_1\tau p_{\mathrm{T}}20}$ events | 22069 | 9684 | 189 | 7476 | 805080 | 6 | 17452 | 7315 | 284 | 7562 | 590115 | 6 |
| $N_{1\ell\_trigg\_1\tau p_{\mathrm{T}}20\_1\tau}$ events | 3049 | 506 | 1 | 136 | 23 | 0 | 1681 | 190 | 0 | 25 | 7 | 0 |
| $N_{1\ell\_trigg\_2\tau p_{\mathrm{T}}20}$ events ($1\ell 2\tau$ events) | 2670 | 363 | 0 | 104 | 21 | 0 | 1681 | 190 | 0 | 25 | 7 | 0 |
| $N_{1\tau}$ events | 37486 | 22362 | 0 | 0 | 1182328 | 1301502 | 27692 | 14793 | 0 | 0 | 803086 | 763144 |
| $N_{1\tau p_{\mathrm{T}}20}$ events | 35331 | 20113 | 0 | 0 | 1090836 | 1124963 | 27692 | 14793 | 0 | 0 | 803086 | 763144 |
| $N_{1\tau p_{\mathrm{T}}20\_1\tau}$ events | 6960 | 2038 | 0 | 0 | 71619 | 0 | 3724 | 916 | 0 | 0 | 33355 | 0 |
| $N_{2\tau p_{\mathrm{T}}20}$ events | 6118 | 1637 | 0 | 0 | 60823 | 0 | 3724 | 916 | 0 | 0 | 33355 | 0 |
| $N_{2\tau p_{\mathrm{T}}20\_1\tau}$ events | 520 | 79 | 0 | 0 | 0 | 0 | 231 | 28 | 0 | 0 | 0 | 0 |
| $N_{3\tau p_{\mathrm{T}}20}$ events ($3\tau$ events) | 434 | 54 | 0 | 0 | 0 | 0 | 231 | 28 | 0 | 0 | 0 | 0 |

*Table 6*