Heavy neutrino analysis update

USTC group meeting

Mário José Sousa

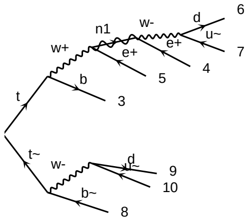
University of Science and Technology of China

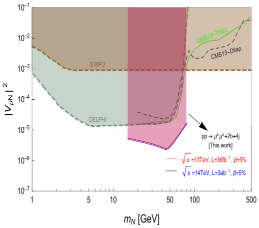
September 8th, 2021

Outline

- Introduction
- Event selection and analysis strategy
- Definition of signal and control regions
- N1 and N2 composition and regions
- Multivariate analysis
- Summary and next steps

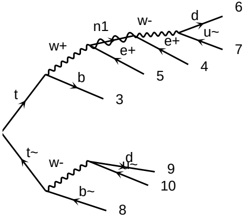
Motivation





- Neutrinos are not massless as it was predicted by the standard model.
- The seesaw mechanism theory might explain the origin of neutrinos masses
- Type-I seesaw is the simplest which include the addition of 3 heavy neutrinos.
- Neutrino-less double beta decay: a lepton number violation of $|\Delta L| = 2$.
- ATLAS and CMS have searches for heavy neutrinos produced with W boson.
- LHC is a top factory and this analysis searches for the heavy neutral leptons in association with the $t\bar{t}$ process.
- Reference theory paper: **link**
- Previous presentation can be found **here**.

Signal characterization



- 1 top quark decays as the Standard Model and the other via a HNL (BSM).
- SM top can decay hadronically (2L) or semileptonically (3L).
- Considering only N1 (coupling to electrons) and N2 (coupling to muons):
  - no mixing ($e - \mu$ final state).*
- Considered only $N \to \ell q \bar{q}'$ semileptonic decay.
- Mass points considered: 15, 25, 35, 40, 45, 50, 55, 60, 70 and 75 GeV.
- Handling particle combinations with the help of m_t and m_W using a χ^2 minimization approach.
- For low HNL masses, W from HNL decay will be off-shell and one of its decaying quarks will be very very soft.
  - An alternative category considers a single jet resulting from this W decay.

Object definition (Top framework)

Electrons

- $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.47$ excluding $[1.37, 1.52]$.
- Identification: TightLH.
- Isolation: Gradient $\rightarrow$ PLVTight
- 2015: e24_lhmedium_L1EM20VH, e60_lhmedium, e120_lhloose
- 2016-18: e26_lhtight_nod0_ivarloose, e60_lhmedium_nod0, e140_lhloose_nod0

Muons

- $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$
- Identification: Medium.
- Isolation: FCTight_FixedRad $\rightarrow$ PLVTight
- 2015: mu20_iloose_L1MU15, mu50
- 2016-8: mu26_ivarmedium, mu50

Hadronically decaying taus

- $p_{\mathrm{T}} > 20$ GeV

Jets

- $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 4.5$

B-tagging

- CDI: 2020-21-13TeV-MC16-CDI-2021-04-16_v1.root
- DL1r with b-tagging efficiency WP 77%.

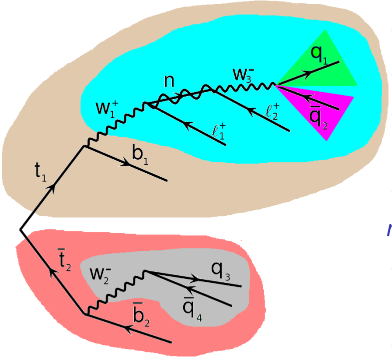
Preselection

- At least an electron or a muon pair.
- At least 3 jets.
- At least 2 b-tagged jets @77% WP.
  - Information for 60% and 70% also saved.

Further selection

- In addition to the top framework preselection:
  - The leading lepton with $p_{\mathrm{T}} > 25$ GeV.
  - Central ($|\eta| < 2.5$) jet selection with $p_{\mathrm{T}} > 20$ GeV.
  - Forward ($|\eta| > 2.5$) jet selection with $p_{\mathrm{T}} > 35$ GeV, passing the forward jet vertex tagger (fJVT).
  - 2 b-tagged central jets passing the 70% efficiency working point with $p_{\mathrm{T}} > 25$ GeV.
  - 2 reconstructed leptons (electrons or muons) with the same flavour.
    - Opposite sign lepton allowed in the most recent strategy.
- Minimum number of particles to progress in the analysis:
  - $\#$non-b-jets ≥ 1
  - 1j: $2 \times (\#\ell - 2)+$ $\#$non-b-jets ≥ 3
  - 2j: $2 \times (\#\ell - 2)+$ $\#$non-b-jets ≥ 4
- Separation between leptons: $0.4 < \Delta R(\ell_1, \ell_2) < 2.5$.

Analysis strategy



Default / 2L

- $q_1 \to j_1$, $\bar{q}_2 \to j_2$
- $q_3 \to j_3$, $\bar{q}_4 \to j_4$
- Top with BSM decay
  $m(t_1) \equiv m(b_1\, \ell_1\, \ell_2\, j_1\, j_2)$
- W with BSM decay
  $m(W_1) \equiv m(\ell_1\, \ell_2\, j_1\, j_2)$
- Top with SM decay
  $m(t_2) \equiv m(b_2\, j_3\, j_4)$
- W with SM decay
  $m(W_2) \equiv (j_3\, j_4)$

3L

- $q_3 \to \ell_3$, $\bar{q}_4 \to \nu$
  $m(t_2) \equiv m(b_2\, \ell_3\, \nu)$
  $m(W_2) \equiv (\ell_3\, \nu)$

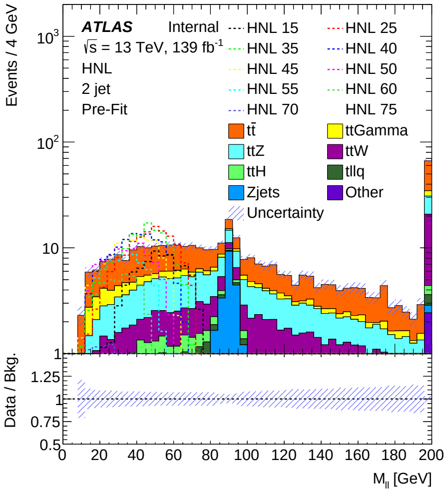
Decision:

- $\chi^2 = \sum_i^4 \frac{(m_i - M)^2}{\sigma_i}$
  - $M_t = 172.76$ GeV
  - $M_W = 80.379$ GeV
- $\chi^2_{2L} < \chi^2_{3L} \Rightarrow$ 2 lepton channel.
- $\chi^2_{2L} > \chi^2_{3L} \Rightarrow$ 3 lepton channel.

Alternative category:

- $\bar{q}_2 \to \not{j}_2$ and same as before.
- 2j events $\subset$ 1j events

Signal and control regions



- W_1 decays to two same sign leptons and W_3 and the invariant mass of the dilepton system invariant mass should not be larger than the mass of the W boson.
- The signal region is defined by $m_{\ell\ell} < 80$ GeV (lt80).
- Two control regions are defined:
  - in the Z mass window: $80 < m_{\ell\ell} < 100$ GeV (mZ)
  - in the high mLL region: $m_{\ell\ell} > 100$ GeV (ht100).
- $t\bar{t}$ is the leading background for this analysis but same sign leptons results from
  - Photon conversion
  - Charge flip of one of leptons
  - Lepton from semileptonic decay from heavy hadron
- Background composition depends on the lepton channel:
  - Splitting $e^{\pm}e^{\pm}$ from $\mu^{\pm}\mu^{\pm}$
- OS control region: opposite sign leptons with $m_{\ell\ell} < 80$ GeV.

Event based in the IFF classes

- Considering only events with at least 1 prompt lepton.
- Prompt: both leptons are prompt.
- Decay: one lepton results from decay.
- Conv: one lepton results from photon conversion.
- Qflip: one lepton results from charge flip.
- Other: anything else (includes no prompt lepton).

Merging backgrounds:

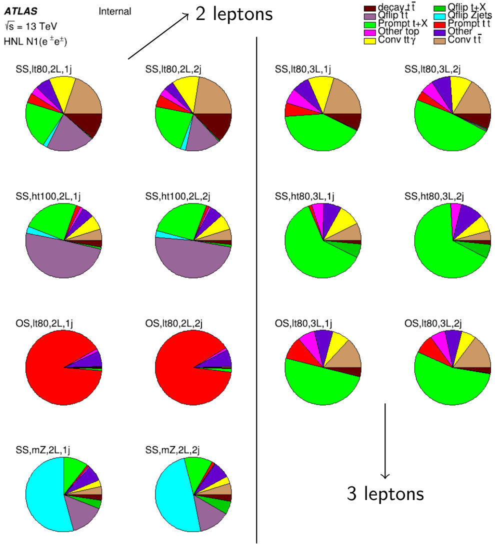
$t\bar{t}$ split into lepton origins, if not negligible.

$t + X$ ttZ, ttW, ttH, ttWW, ttWZ, tllq, 3t, 4t.

Other Top Same as above not considered (other lepton origins) + single top (all channels).

Other : W, Z, VV, VVV

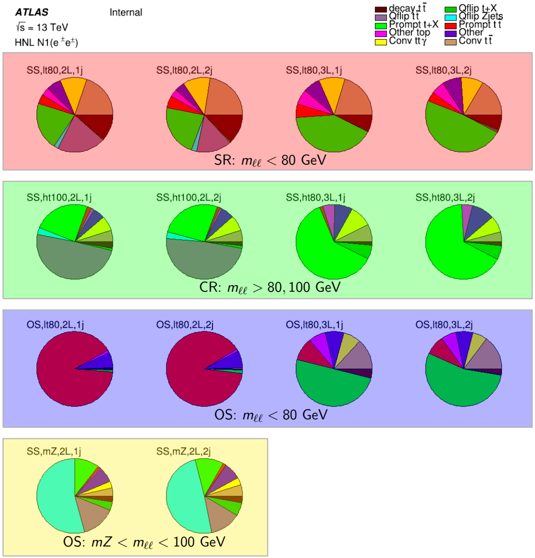
Composition for background in diverse regions: electron channel



Food for thought

1. Use mZ region to normalize Z+jets Qflip or remove (negligible in SR).
2. Use 2L OS to normalize $t\bar{t}$-Prompt in other regions.
3. Use 2L SS ht100 to normalise $t\bar{t}$-Qflip 2L SR.
4. Use 3L SS ht80 to normalize t+X-Prompt in other regions.
5. $t\bar{t}$-Conv and $t\bar{t}\gamma$-Conv normalized (together) from SS 3L ht80 (after 1 and 4).
6. No ideas for $t\bar{t}$-Decay background (see composition in backup).

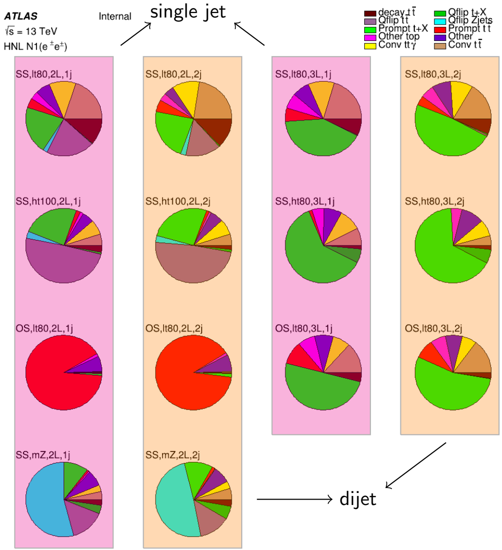
Composition for background in diverse regions: electron channel



Food for thought

1. Use mZ region to normalize Z+jets Qflip or remove (negligible in SR).
2. Use 2L OS to normalize $t\bar{t}$-Prompt in other regions.
3. Use 2L SS ht100 to normalise $t\bar{t}$-Qflip 2L SR.
4. Use 3L SS ht80 to normalize t+X-Prompt in other regions.
5. $t\bar{t}$-Conv and $t\bar{t}\gamma$-Conv normalized (together) from SS 3L ht80 (after 1 and 4).
6. No ideas for $t\bar{t}$-Decay background (see composition in backup).

Composition for background in diverse regions: electron channel

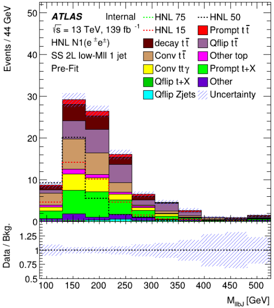


Food for thought

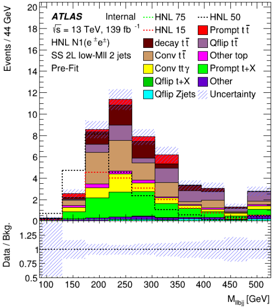
1. Use mZ region to normalize Z+jets Qflip or remove (negligible in SR).
2. Use 2L OS to normalize $t\bar{t}$-Prompt in other regions.
3. Use 2L SS ht100 to normalise $t\bar{t}$-Qflip 2L SR.
4. Use 3L SS ht80 to normalize t+X-Prompt in other regions.
5. $t\bar{t}$-Conv and $t\bar{t}\gamma$-Conv normalized (together) from SS 3L ht80 (after 1 and 4).
6. No ideas for $t\bar{t}$-Decay background (see composition in backup).

Signal regions: electron channel

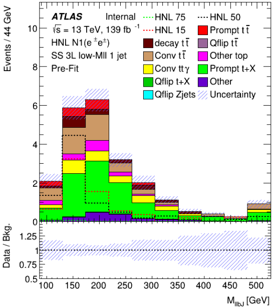
2L 1j



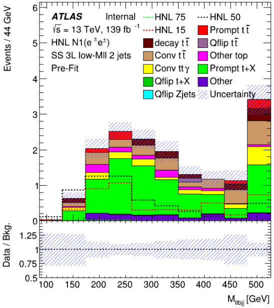
2L 2j



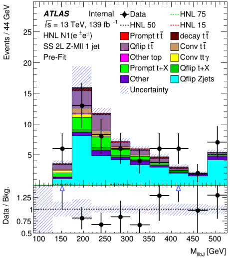
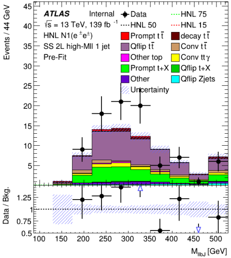
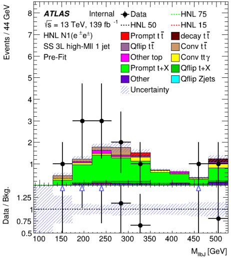
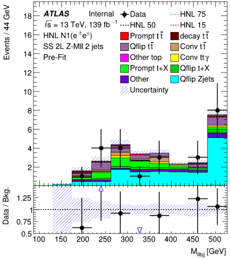
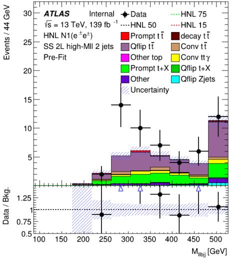
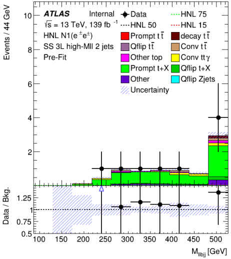
3L 1j



3L 2j

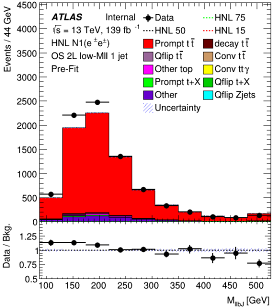


Same sign control regions: electron channel

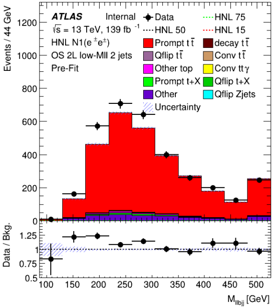
	2L $m_{\ell\ell} \in m_Z$	2L $m_{\ell\ell} > 100$ GeV	3L $m_{\ell\ell} > 80$ GeV
1j			
2j			

Opposite sign control regions: electron channel

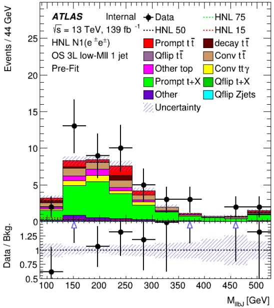
2L 1j



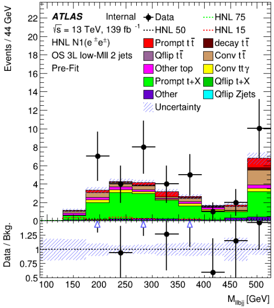
2L 2j



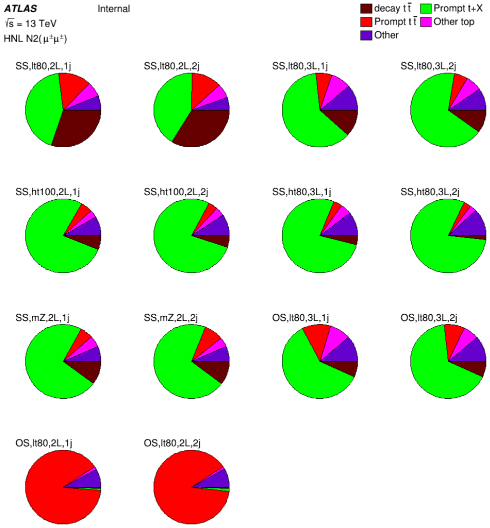
3L 1j



3L 2j



Composition for background in diverse regions: muon channel

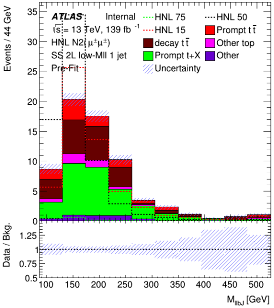


Food for thought

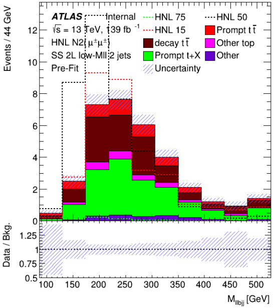
1. Use 2L OS to normalize $t\bar{t}$-Prompt in other regions.
2. Use 2L/3L SS ht100/ht80 to normalize t+X-Prompt in other regions.
3. $t\bar{t}$-Decay background normalized in mZ and OS 3L control regions

Signal regions: muon channel

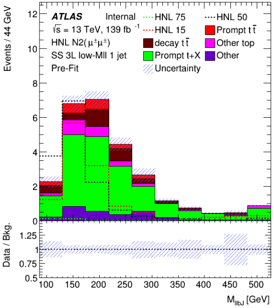
2L 1j



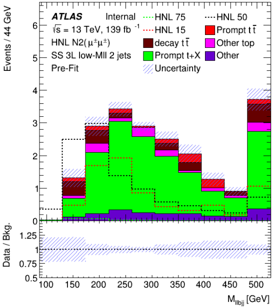
2L 2j



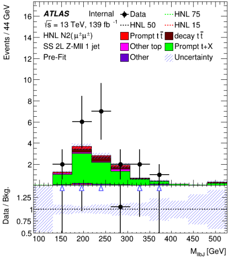
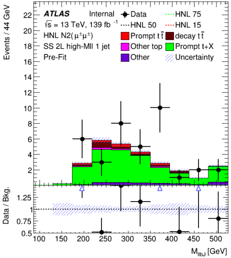
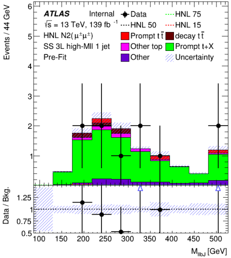
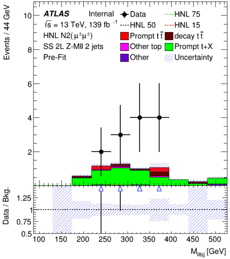
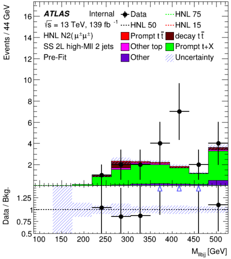
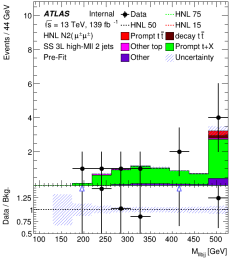
3L 1j



3L 2j

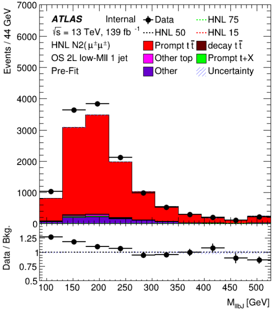


Same sign control regions: muon channel

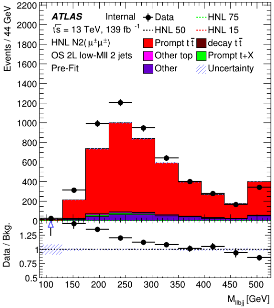
	2L $m_{\ell\ell} \in m_Z$	2L $m_{\ell\ell} > 100$ GeV	3L $m_{\ell\ell} > 80$ GeV
1j			
2j			

Opposite sign control regions: muon channel

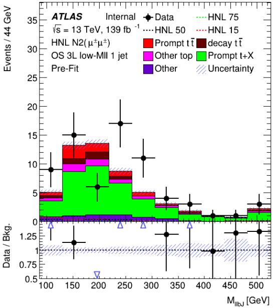
2L 1j



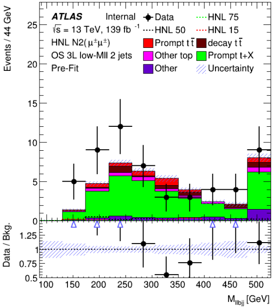
2L 2j



3L 1j

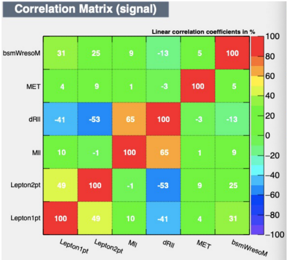


3L 2j

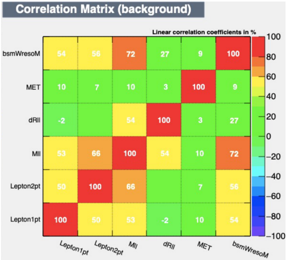


Multivariate analysis

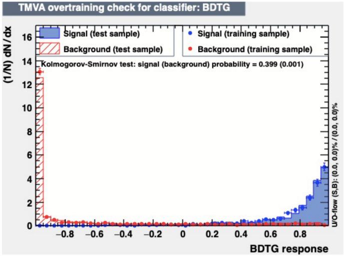
Signal



Background



Overtraining check



Variable importance

Rank	Variable	Importance
1	$\Delta R(\ell,\ell)$	24.77 %
2	$\not{E}_T$	17.61 %
3	$m_{\ell\ell jj}$	16.77 %
4	$p_T^{\ell 1}$	14.23 %
5	$m_{\ell\ell}$	13.33 %
6	$p_T^{\ell 2}$	13.30 %

-
- I will start to be more envolved in the near future.
- Starting training stage with 17 variables in 1 signal region 2j.
- Optimization to reduce variables down to 6

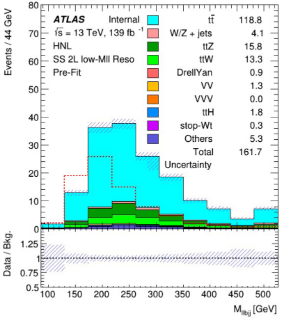
Multivariate analysis

Cut-based

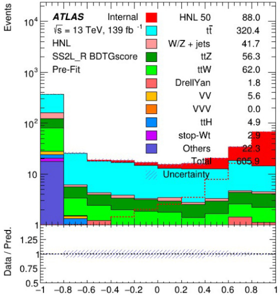
- $12 < m_{\ell\ell} < 80$ GeV
- $0.4 < \Delta R(\ell\ell) < 2.5$
- Expected significance: 6.5

MVA-based

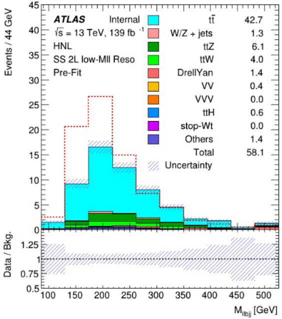
- $m_{\ell\ell} > 12$ GeV
- BDTGscore >0.2
- Expected significance: 8.8



Cut-based



BDTG Score



MVA-based

Summary

- The existence of heavy neutral leptons would prove the seesaw mechanism theory to explain the neutrino masses.
- This is the first analysis probing this theory in a $t\bar{t}$ associated production.
- Analysis strategy close to final.
- Preliminary MVA studies show an improvement in sensibility from 6.5 to 8.8 in one region.

Next steps:

- Signal production of the fully leptonic decay of HNL.
  - Yields doubles in 3L signal region at 50 GeV mass point.
- Documentation
- Finalise analysis strategy
- Systematics
- MVA studies in other regions

BACKUP

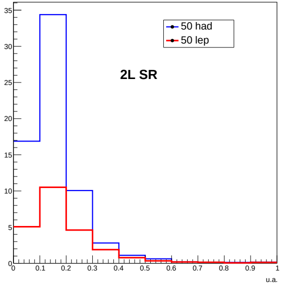
Backup: HNL full leptonic decay

- Additional signal is now ready and present in the top+X eos folder.
- Differences comes from using the fully leptonic decay of the HNL.

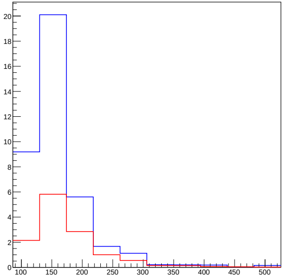
	SS 2L low-Mll 1 jet	SS 2L low-Mll 2 jets	SS 3L low-Mll 1 jet	SS 3L low-Mll 2 jets	SS 2L high-Mll 1 jet	SS 2L high-Mll 2 jets
HNL 50	66.4	39.3	14.3	10.5	0.00	0.00
HNL 50 lep	23.6	12.1	16.2	12.9	0.06	0.03

HNL 50	66.4	39.3	14.3	10.5	0.00	0.00	0.03	0.02	0.00	-0.05	2.00	1.41	4.39	2.16
HNL 50 lep	23.6	12.1	16.2	12.9	0.06	0.03	0.26	0.18	0.15	0.05	7.90	6.37	8.59	4.07

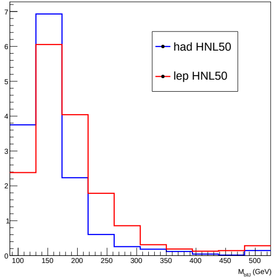
2L 1j MU



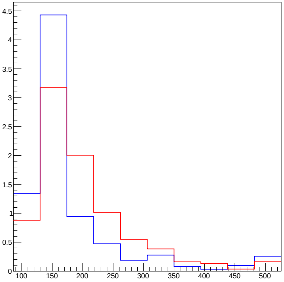
2L 1j EL



3L 1j MU



3L 1j EL

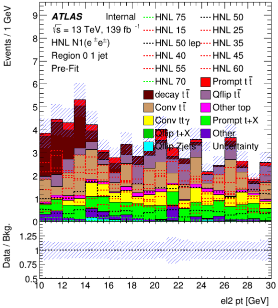


For muons

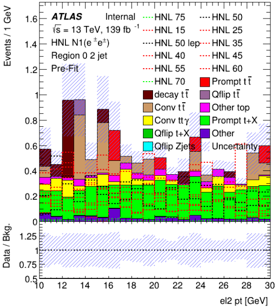
- In 2L SR, signal is increased by 35.5% (30.8%) in 1j (2j) regions.
- In 3L SR, signal is increased by 113% (123%) in 1j (2j) regions.
- Recommendation: request other HNL decaying leptonically.

Backup: idea for $t\bar{t}$ decay

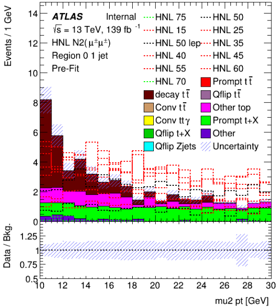
Electron 1j



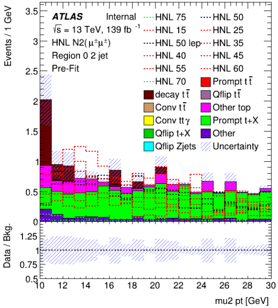
Electron 2j



Muon 1j

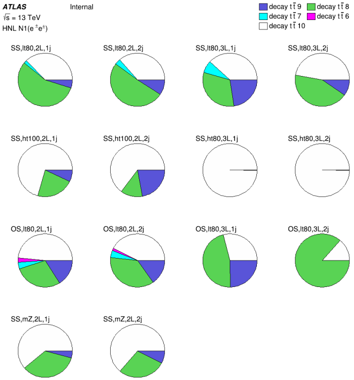


Muon 2j

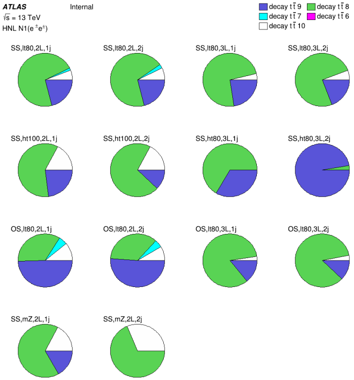


- Events in which subleading (non-prompt) lepton has very low-p_{T} concentrate the cases where this lepton results from a hadronic decay.
- Increasing the lepton threshold ($>$ 15 GeV?) might reduce this background to acceptable values.

EL



MU



Based on this input, the
(IFF classes):

0 • Unknown
1 • KnownUnknown
2 • IsoElectron
3 • ChargeFlipIsoElectro
4 • PromptMuon
5 • PromptPhotonConv
6 • ElectronFromMuon
7 • TauDecay
8 • BHadronDecay
9 • CHadronDecay
10 • LightFlavorDecay

More details about the le

IFF classes

0 Unknown
1 KnownUnknown
2 IsoElectron
3 ChargeFlipIsoElectron
4 PromptMuon
5 PromptPhotonConversion
6 ElectronFromMuon
7 TauDecay
8 BHadronDecay
9 CHadronDecay
10 LightHadronDecay

- 2,4 as prompt
- 6,7,8,9,10 as decay
- 3 as charge flip
- 5 as photon conversion
- 0 and 1 as other
- -1 as not filled

More information **here**

The KnownUnknown category refers to leptons which can (in principle) be classified, but the tool fails with the classification due to missing information.