Neural Networks applied to Signal–Background discrimination in Flavoured Dark Matter searches with the ATLAS detector

Jordi Barberà Donat

Tutors:

Imma Riu
Álvaro López

Contents

Contents

I. 1. Context and motivation

- ➢ From a cosmological and astrophysical point of view, the existence of a non-baryonic and weakly interacting —electromagnetically neutral (*dark*)— component of matter is widely accepted

- ➢ Even though dark matter accounts for the 26% of the energy budget of the Universe, its fundamental component —a DM particle— has not yet been observed

- ➢ Cosmologically, DM is currently considered as a main ingredient in the evolution of the Universe and the structure formation

- ➢ In the past few years, the LHC has been studying the possibilities of simple DM models. Recently, the ATLAS Collaboration and, in general, the HEP Community has begun the study of more complex models: introducing QCD-like interactions and different DM particle families

I. 2. Theoretical model

- The study is enclosed in the framework of a 3 particle family DM model, with couplings to the Standard Model connecting different quarks and DM generations
- Coupling is mediated by a coloured scalar ϕ with the same quantum numbers as quarks

$$\mathcal{L}_{\text{dark}} = (i\bar{\chi}\not{\partial}\chi - m_\chi\bar{\chi}\chi) - (i\lambda_{ij}\bar{q}_{L,i}\phi\chi_j + \text{h.c.}) + (D_\mu\phi)^\dagger(D^\mu\phi) - m_\phi^2\phi^\dagger\phi - V(\phi, H)$$

- Main free parameters of the model: **masses** of the mediator and DM particles, **couplings** and **mixing angles**

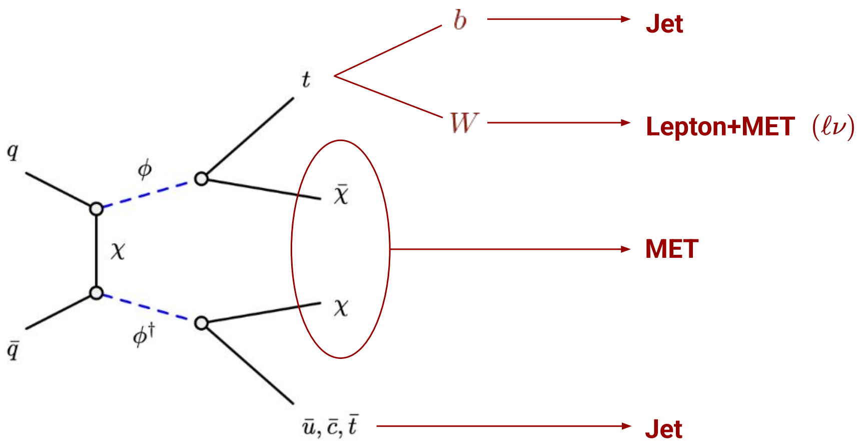
$$\lambda = U_\lambda D_\lambda \quad \text{with} \quad D_\lambda = \text{diag}(D_{\lambda,11}, D_{\lambda,22}, D_{\lambda,33})\ , \quad U_\lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23}e^{-i\delta_{23}} \\ 0 & -s_{23}e^{i\delta_{23}} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{13}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12}e^{-i\delta 12} & 0 \\ -s_{12}e^{i\delta_{12}} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$m_{\chi,ij} = m_\chi(1 + \eta(D_{\lambda,ii})^2 + \mathcal{O}(\lambda^4))\delta_{ij}$$

- DM particles can be either Dirac-type or Majorana-type, and can couple to either right-handed or left-handed quarks

I. 3. Specifications

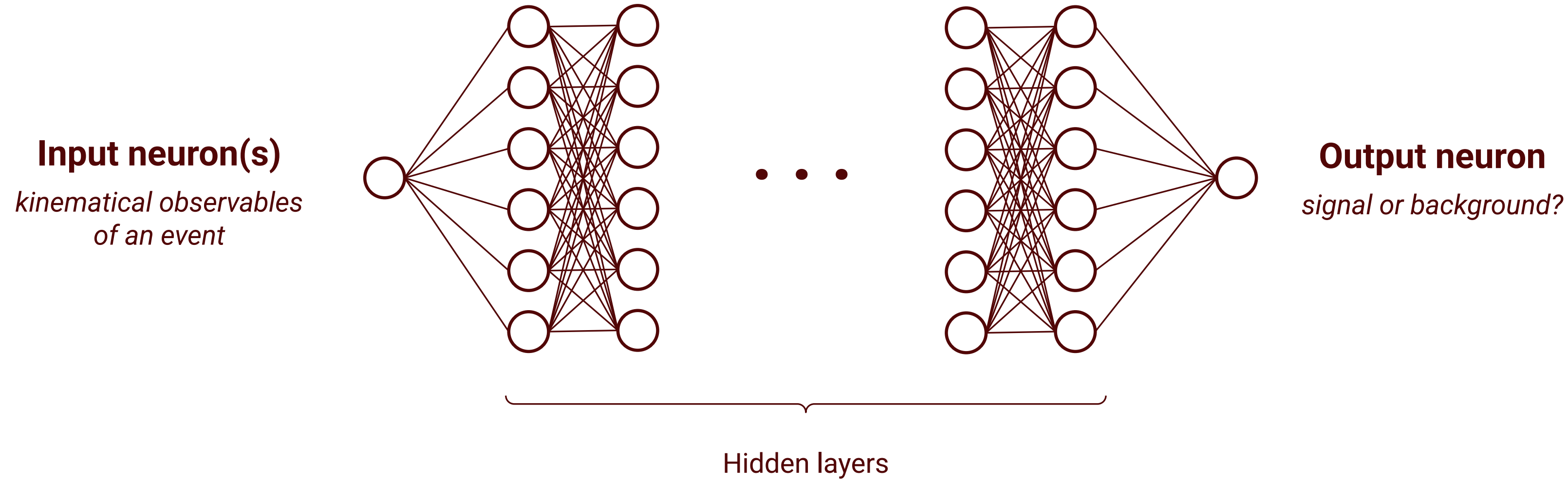
➢ In our study we considered **Dirac-type** DM particles with couplings to **right-handed** quarks, leading to a final state with a top quark that decays to a **1 lepton** state



CONTENTS

II. 1. Neural Network strategy

- Classical strategy on signal-background discrimination consists on applying clever selections, i.e. choosing events with kinematical variables inside some range, so we remove background events keeping as many signal events as possible
- We aim to enhance the discrimination process through the use of a Deep Neural Network, as they are, in general, more capable of capturing correlations between variables



II. 2. DNN hyperparameters

- The neural network has been built-up and trained using the PyTorch library

- The procedure consists on feeding the NN with mixed up signal & background events, so it learns to tag them properly

- The hyperparameters that can be changed in this particular DNN are:

DNN architecture

- Number of hidden layers = 3
- Number of neurons per hidden layer = 32
- Learning rate of the NN = 10^{-5}
  how fast the NN approximates itself to the optimal parameters

Dataset division

- Batch size = 32
  size of subsets in which are divided the total events
- Number of epochs = 8
  number of times all of the events are fed to the NN

CONTENTS

III. 1. Monte Carlo simulations

- Monte Carlo simulations for different DM parameters
- ATLAS detector effect is also being considered in the simulations
- Preset benchmark:

$$\theta_{13} \equiv \theta \qquad D_{\lambda,11} = D_{\lambda,22} \equiv d$$
$$\theta_{12} = \theta_{23} = 0 \qquad D_{\lambda,33} = d + 0.5$$

Background

- t t-bar
- singletop
- z-jets
- w-jets
- ttZ
- Multiboson
- Raretop
- Other (ttH + dijet)

Signals

$$\delta_{12} = \delta_{13} = \delta_{23} = 0$$
$$\theta = 0.1 \qquad \eta = 0$$

III. 2. Signal vs. Background plots. Signal samples

Intermediate DM parameters to study the variable behaviour: $(\theta = 0.1, \quad \eta = 0)$

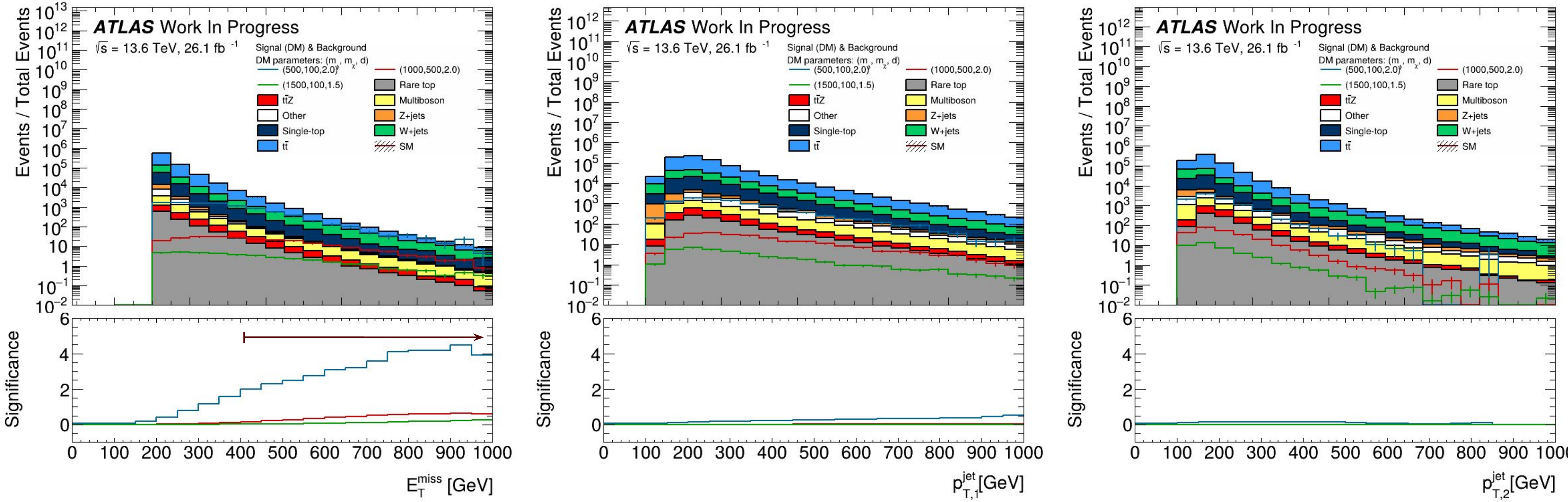
Signal samples	m_ϕ/GeV	m_χ/GeV	d
Sample 1	1500	100	1.5
Sample 2	1000	500	2.0
Sample 3	500	100	2.0

Variables studied :

- Basic kinematics: **MET + p_T**
- Tagged kinematics: **$p_T^{b\text{-jet}} + p_T^{c\text{-jet}}$**
- Angular distances: **ΔR**
- Transverse masses: **m_T**
- Stransverse masses: **m_{T2}**

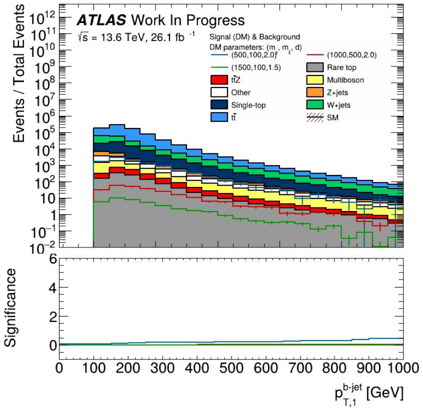
Contents

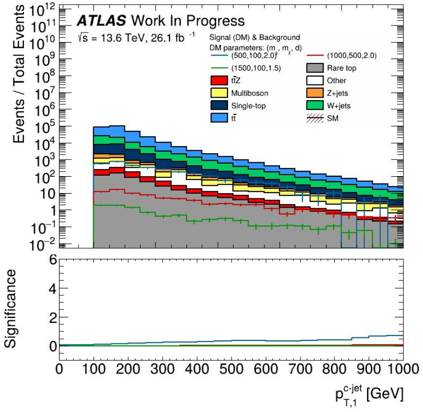
VII. 1. Basic kinematics

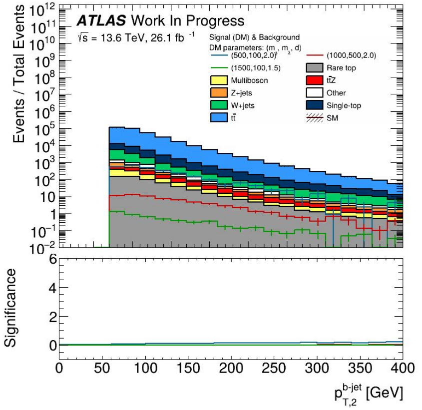


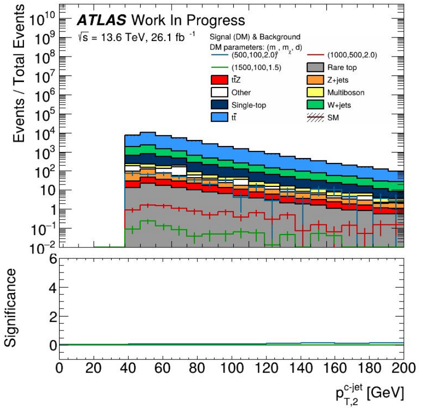
Because of being non-interactive, we expect **high E_T^{miss}** in the presence of DM particles

VII. 2. Tagged kinematics

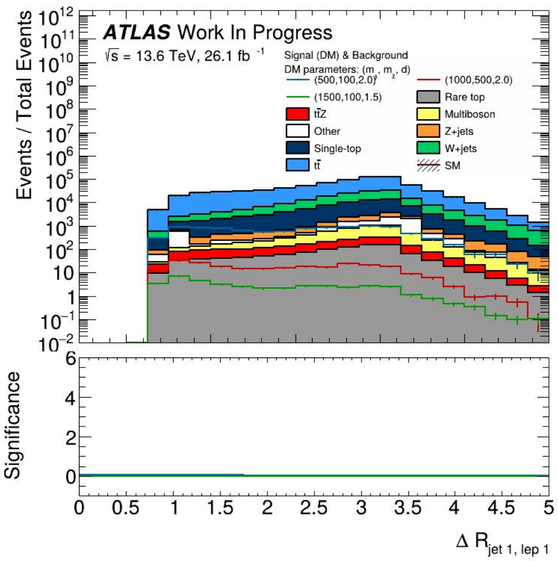


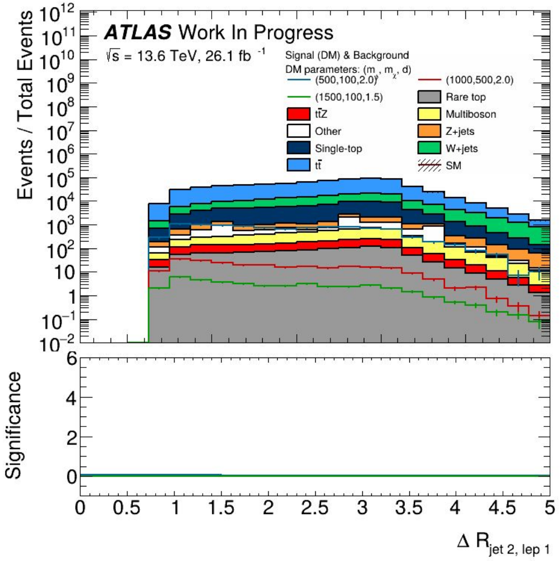


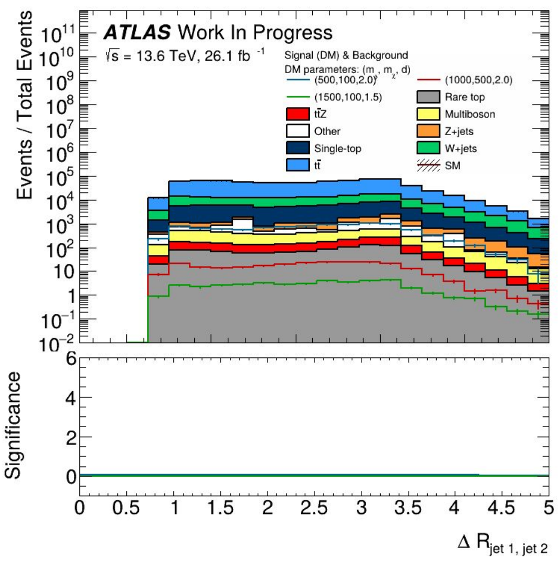




VII. 3. Angular distances

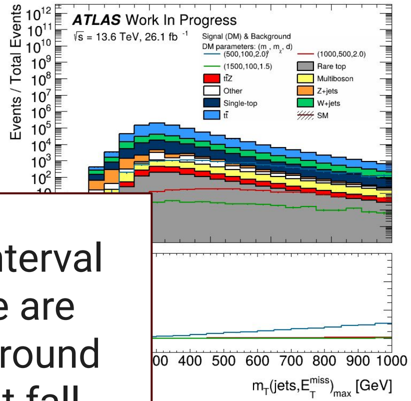


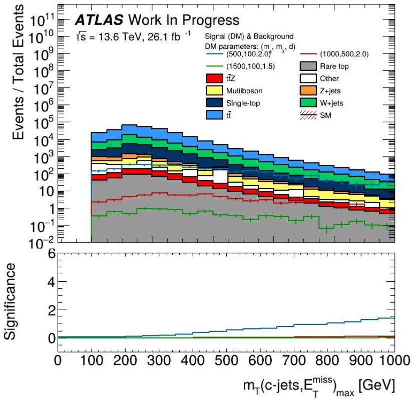


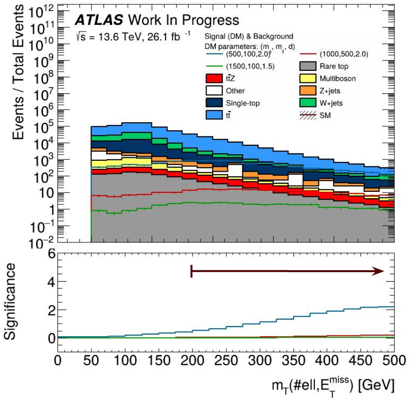


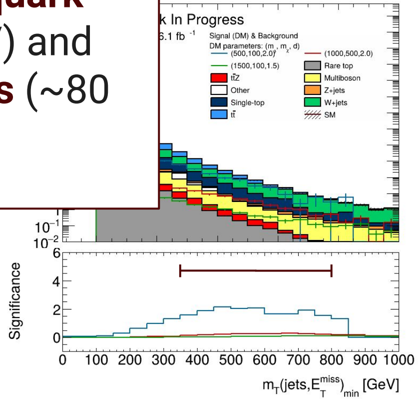
VII. 4. Transverse masses, $m_T = \sqrt{E_T^2 - p_T^2}$

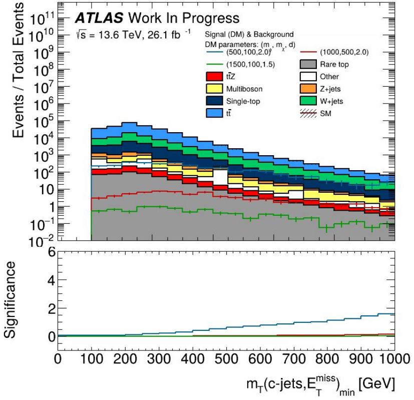
By removing the interval m_T < 200 GeV we are subtracting background contributions that fall around the **top quark mass** (~170 GeV) and the **W boson mass** (~80 GeV)

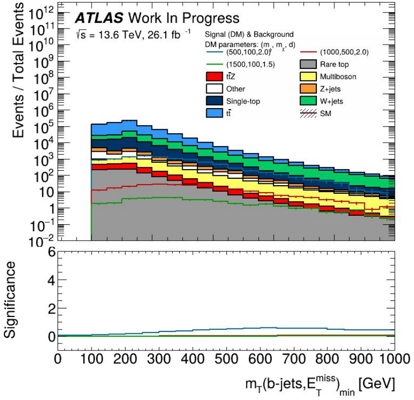




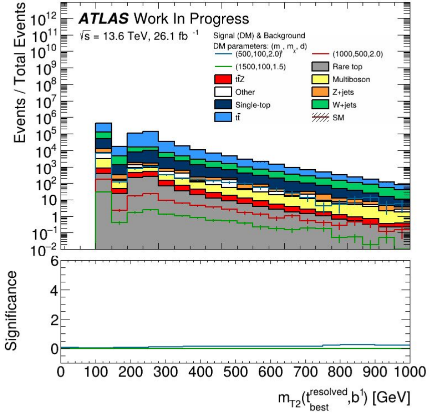


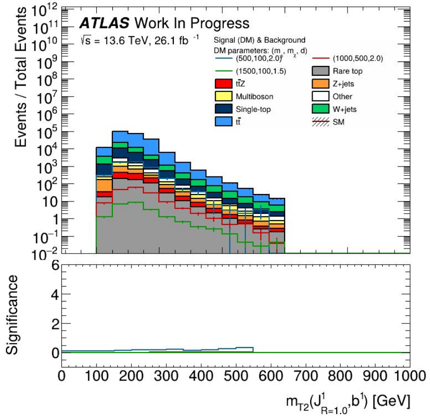


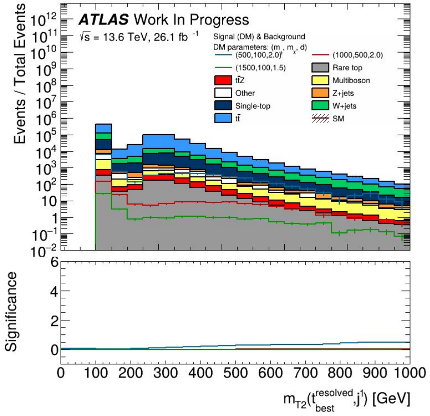


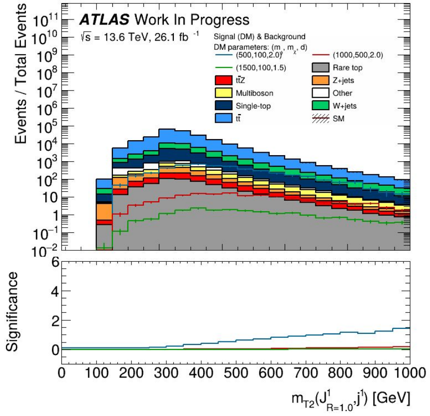


VII. 5. Stransverse masses







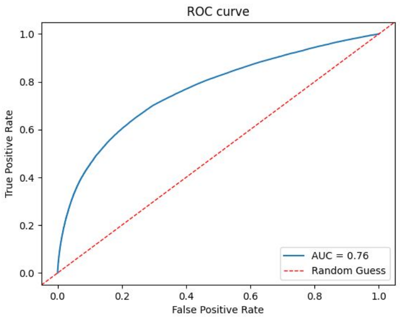


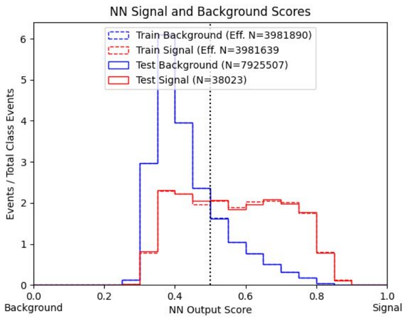
Contents

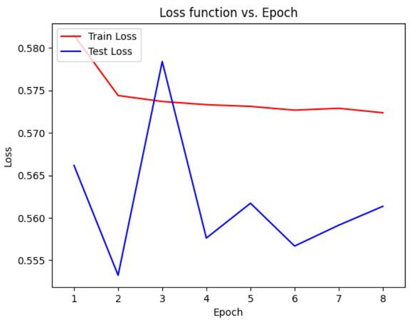
V. 1. Good performance vs. Bad performance

Angular distances

ΔR

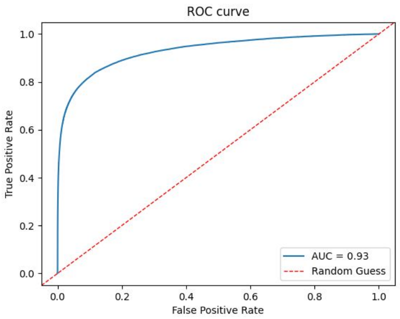


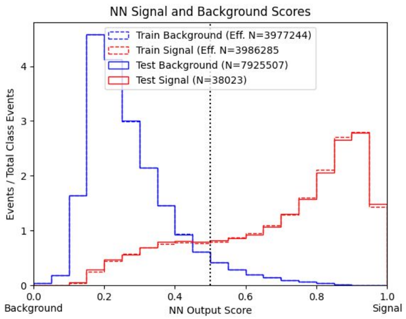


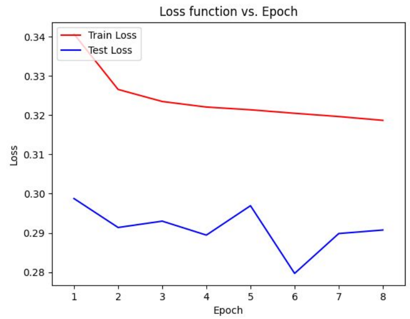


Minimal Most Performative Variables

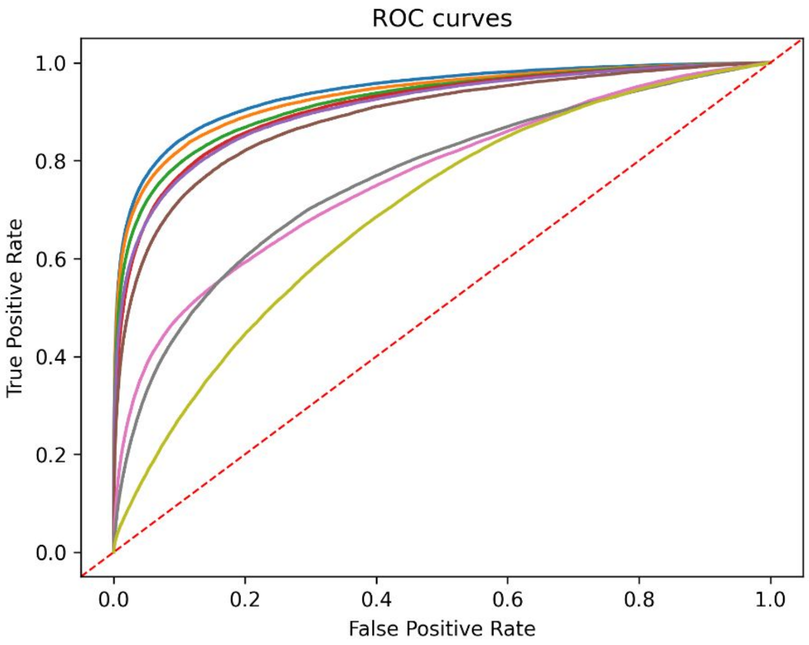
Basic kinematics + m_T

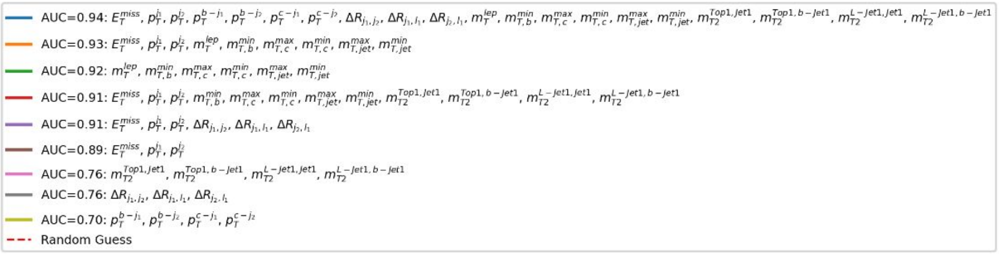






V. 2. Impact of variables in discrimination





1. ALL
2. Minimal Most Performative Variables: Basic kin. + m_T
3. Transverse masses
4. Basic kinematics + Masses (m_T + m_{T2})
5. Basic kinematics + Angular distances
6. Basic kinematics
7. Stransverse masses
8. Angular distances
9. Tagged kinematics

$$\varepsilon = \frac{N_{\text{discriminated}}}{N_{\text{total}}}$$

Contents

VI. Conclusions & Next steps

➢ A DNN-based analysis can be decisive on the signal/background discrimination process

➢ We have been able of extracting conclusions based on the kinematical variables used:

- Stransverse masses, angular distances and tagged kinematics do not contribute significantly to NN performance

- Basic kinematics (AUC = 0.92 with 6 different variables) and transverse masses (AUC = 0.89 with 3 different variables) can achieve a good performance with a low number of variables

- Just with basic kinematics and transverse masses (MMPV), the performance is nearly (AUC = 0.93) the performance achieved with every variable considered (AUC = 0.94)

➢ It would be better to prioritise resolving basic kinematics and transverse masses

➢ Future studies will analyse the effects of combining both cuts and DNN strategies: introducing manual selections of kinematical variables sets and cuts *could* lead to a computational cost saving

Thank you!

Moltes gràcies!

Back-up

Selections & Cuts Signal vs. Background plots

Proposed selections

Significance selections

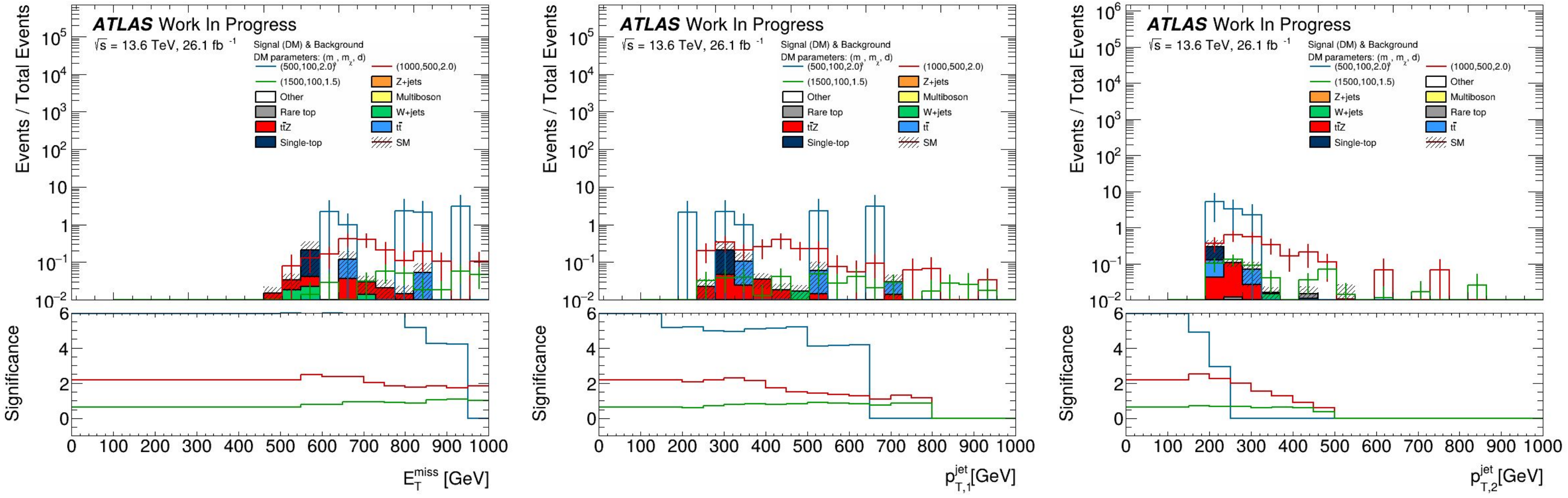
- $N_{lep} = 1$
- $N_{jets} \geq 4$
- $N_{b\text{-}jets} \geq 1$
- $E_T^{miss} \geq 400$ GeV
- 200 GeV $\leq m_T(\text{jets}, E_T^{miss})_{min} \leq$ 650 GeV
- $m_T(\#\text{ell}, E_T^{miss}) \geq 200$ GeV

SR1L1c

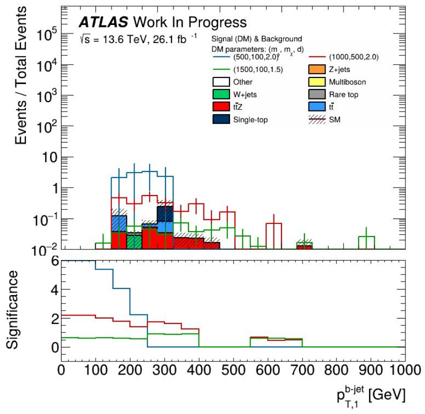
- $N_{jets} \geq 2$
- $N_{c\text{-}jets} = 0$
- $N_{b\text{-}jets} \geq 1$
- $E_T^{miss} \geq 200$ GeV
- $m_T(\text{c-jets}, E_T^{miss})_{max/min} \geq 150$ GeV
- $m_T(\text{b-jets}, E_T^{miss})_{max/min} \geq 200$ GeV
- E_T^{miss} significance ≥ 10
- $m_T(\#\text{ell}, E_T^{miss}) \geq 30$ GeV

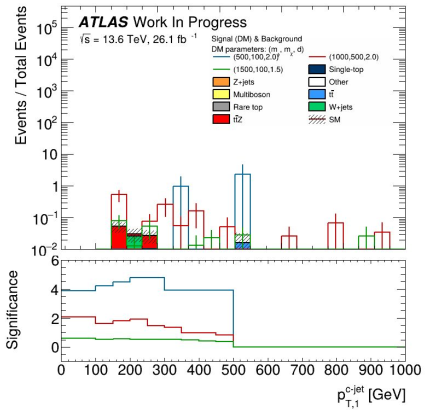
Significance

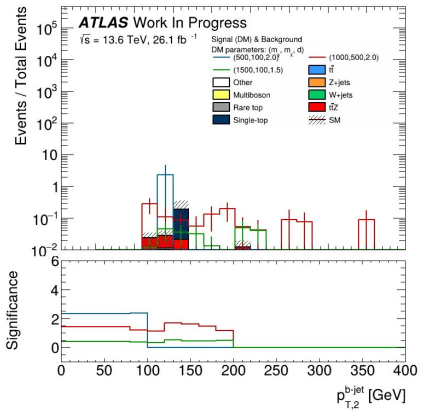
Basic kinematics

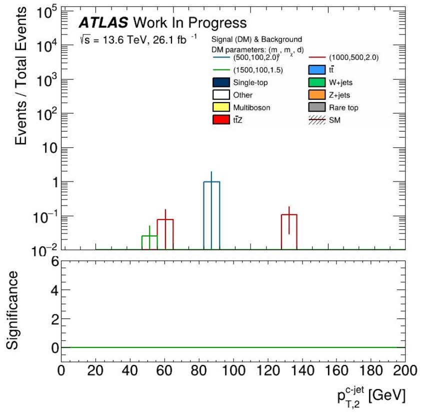


Tagged kinematics

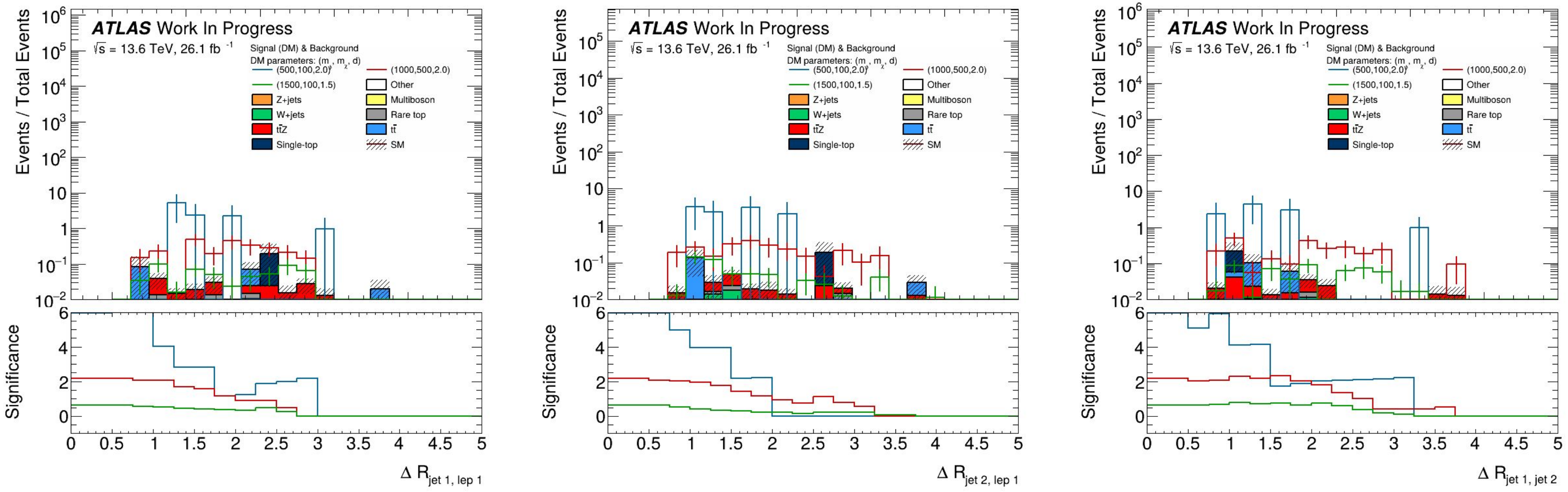




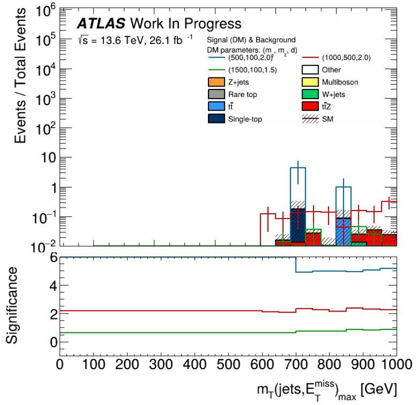


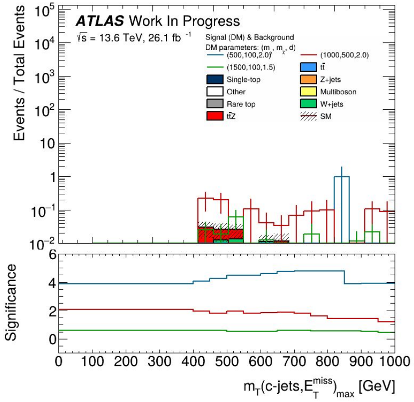


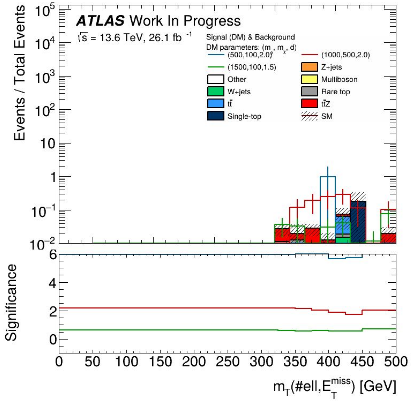
Angular distances

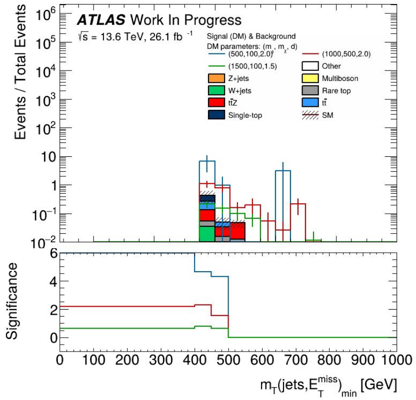


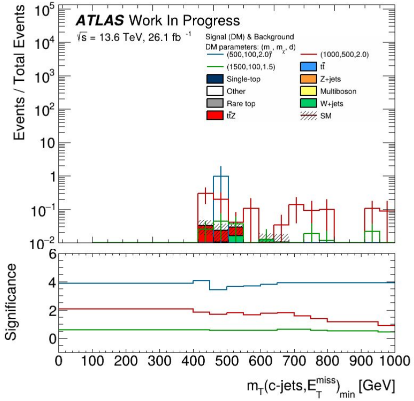
Transverse masses

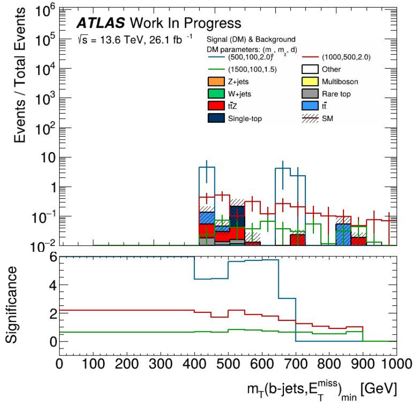




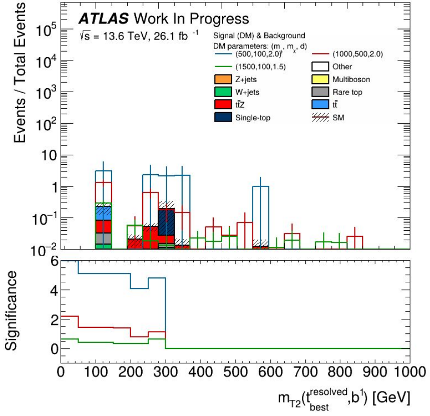


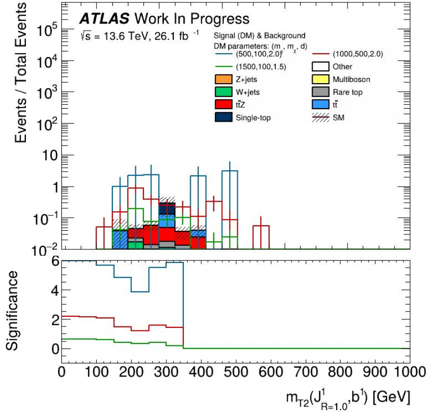


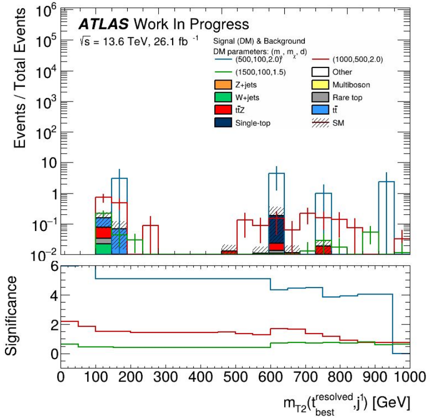


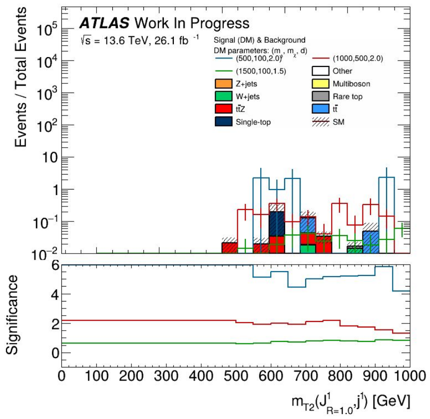


Stransverse masses



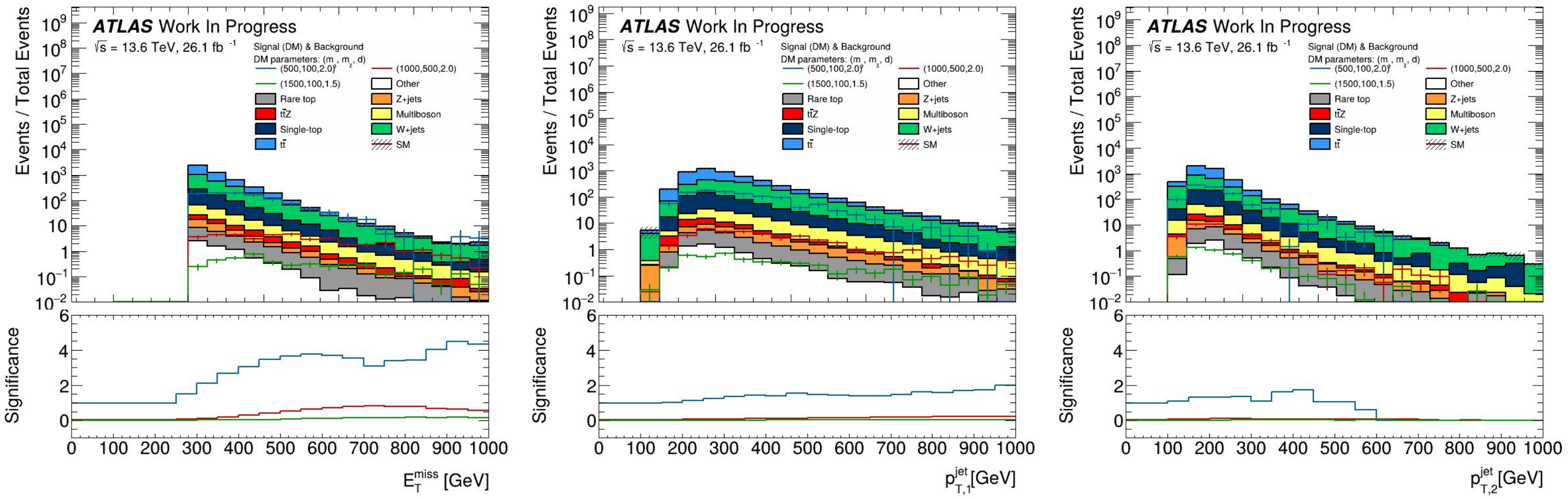




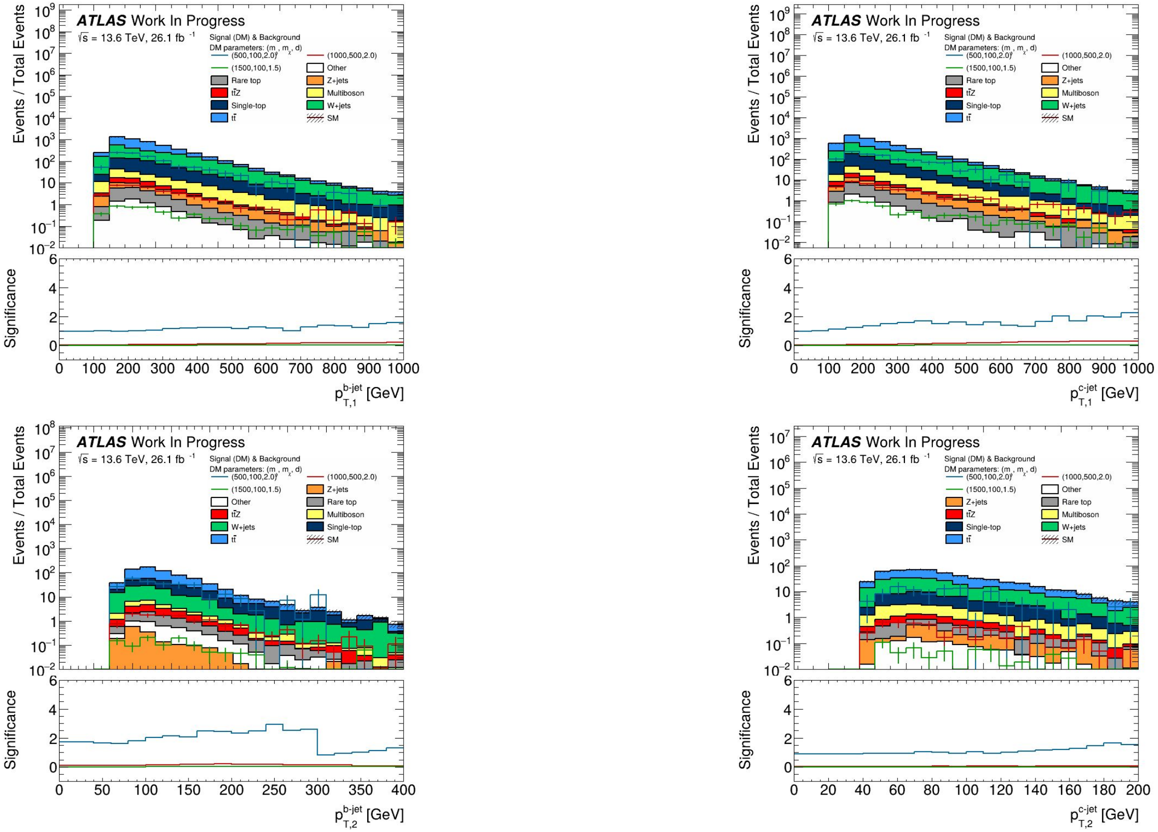


SR1L1c

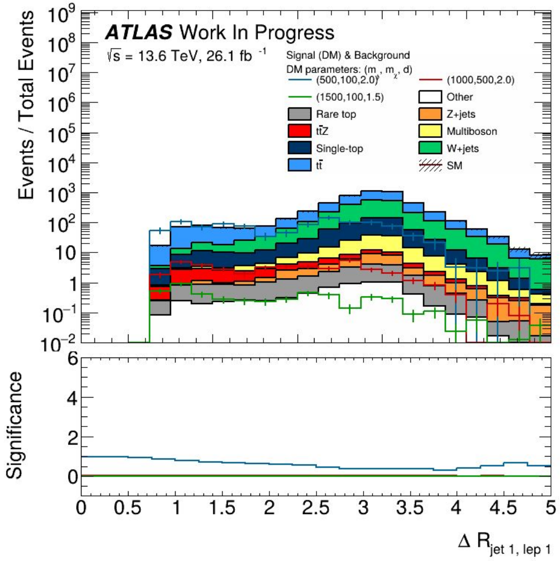
Basic kinematics

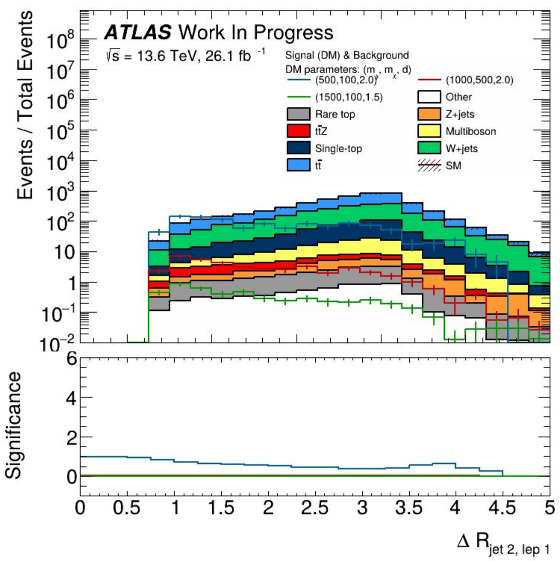


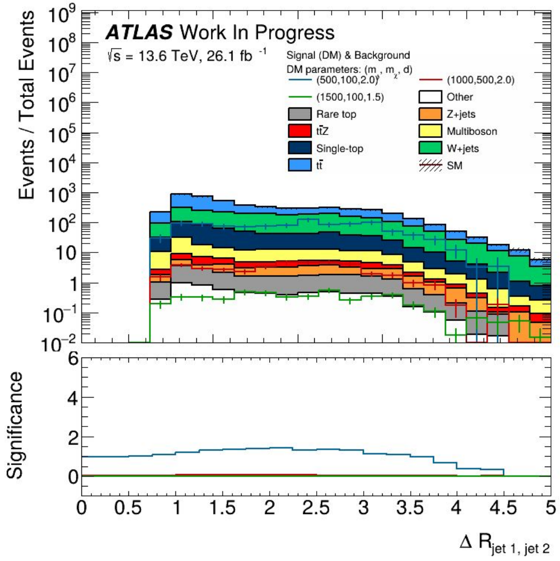
Tagged kinematics



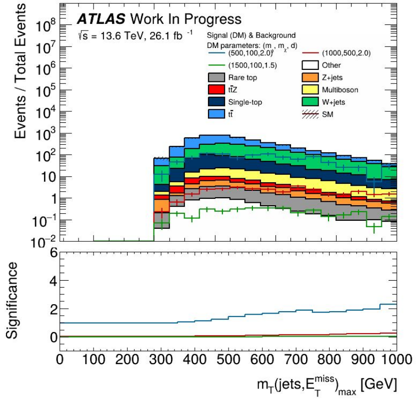
Angular distances

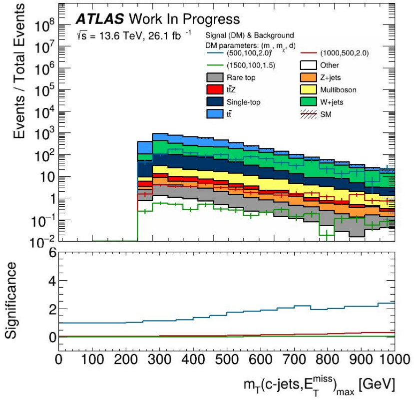


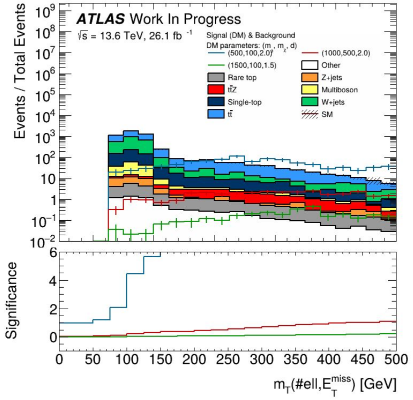


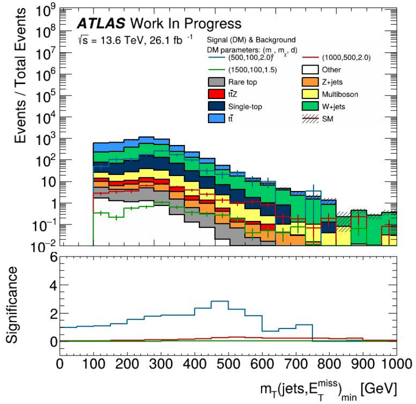


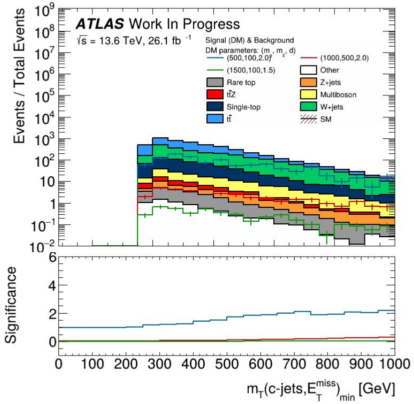
Transverse masses

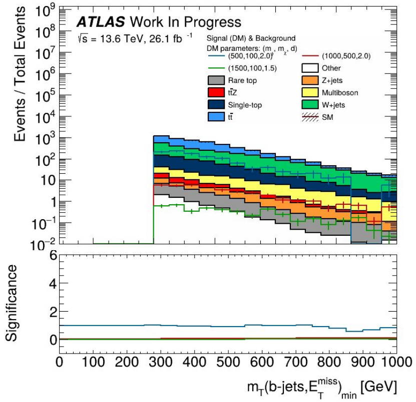




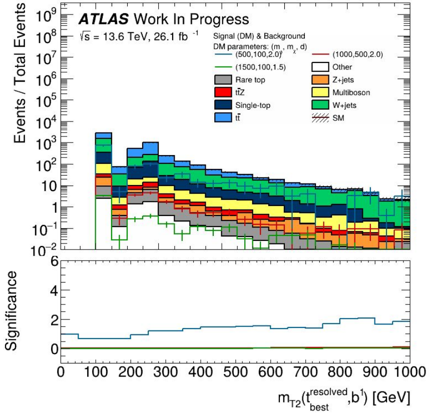


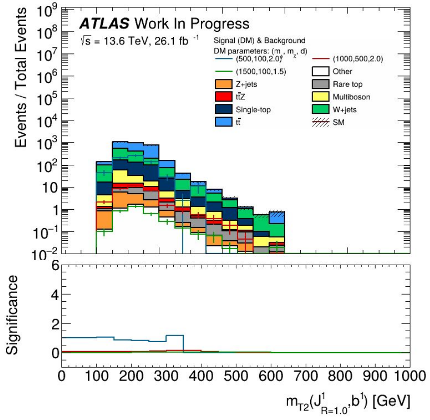


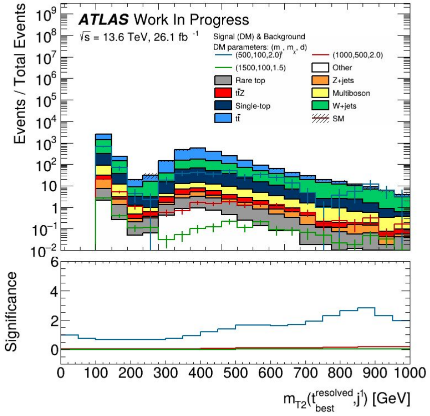


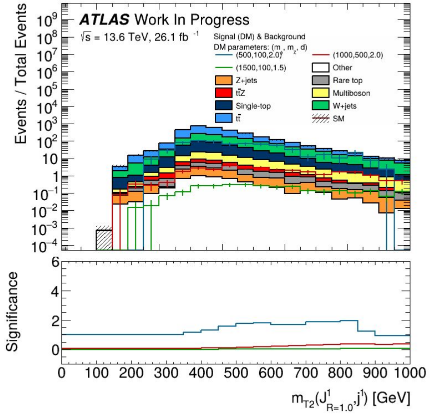


Stransverse masses



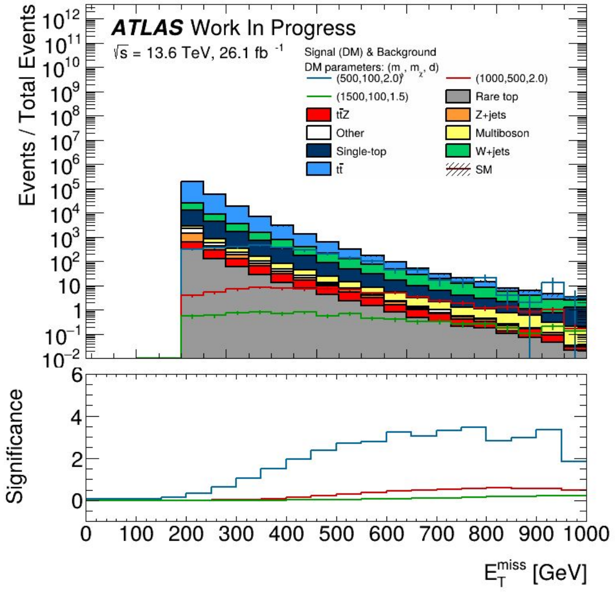


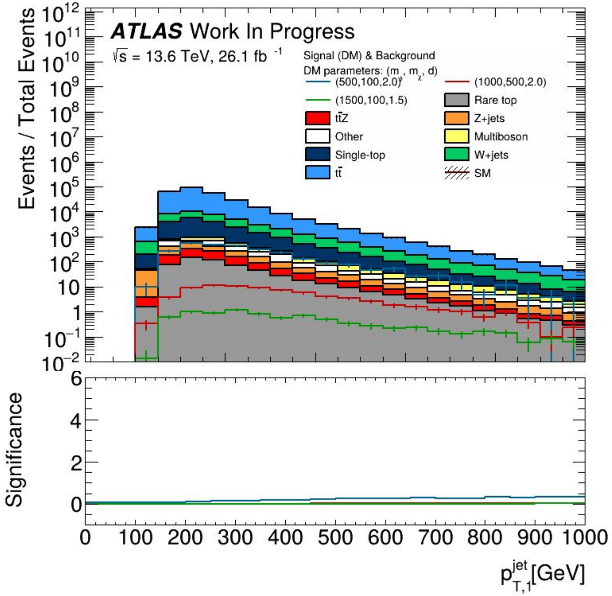


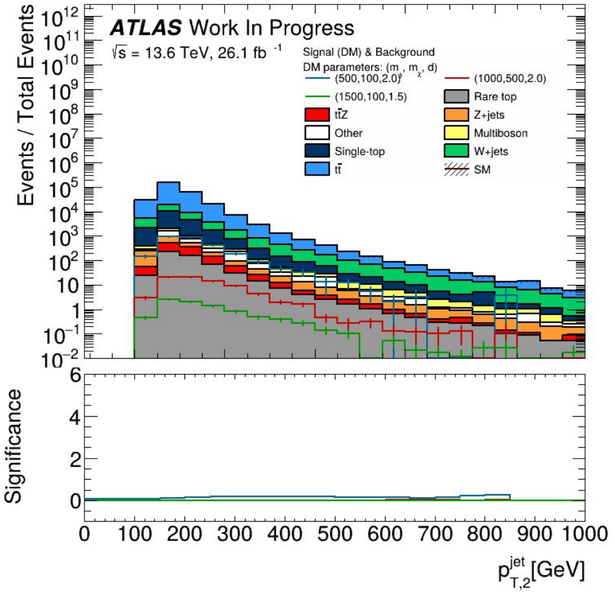


pre-1l-1b-dphimin

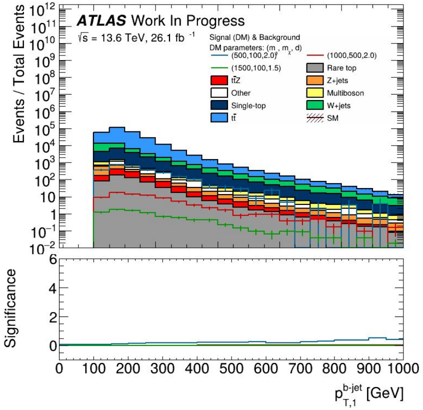
Basic kinematics

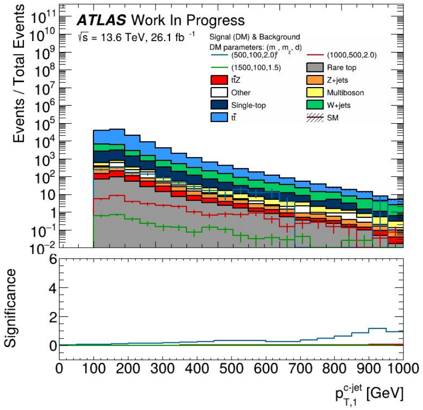


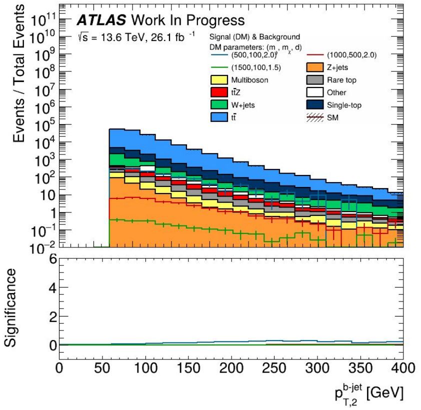


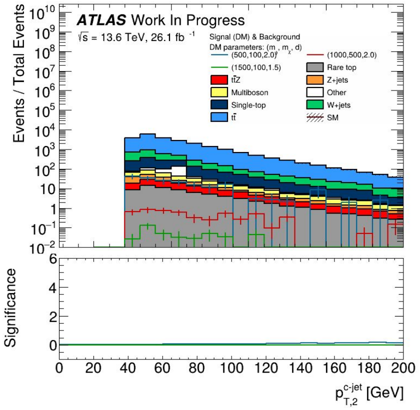


Tagged kinematics

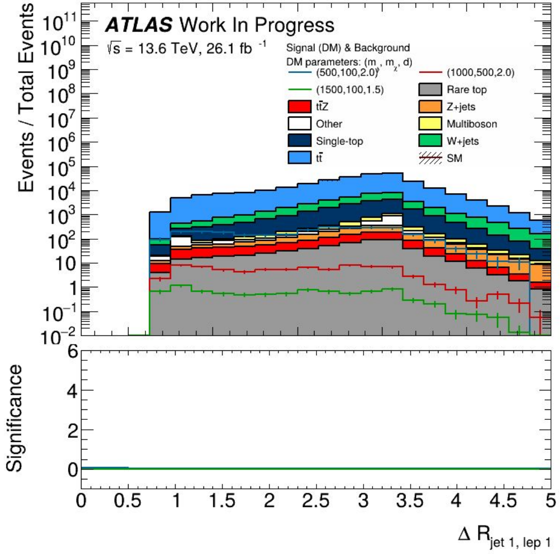


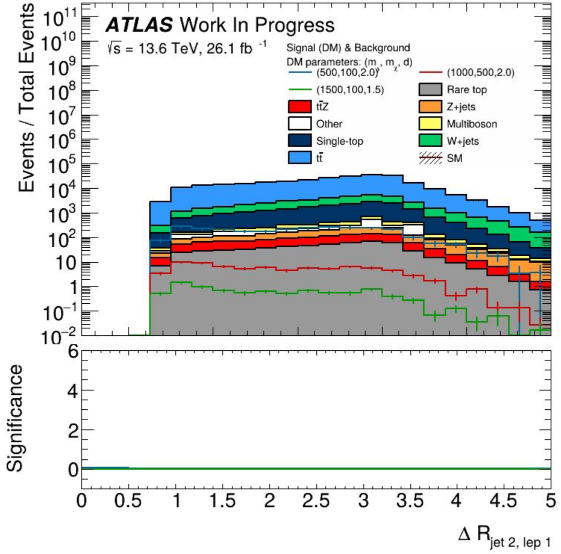


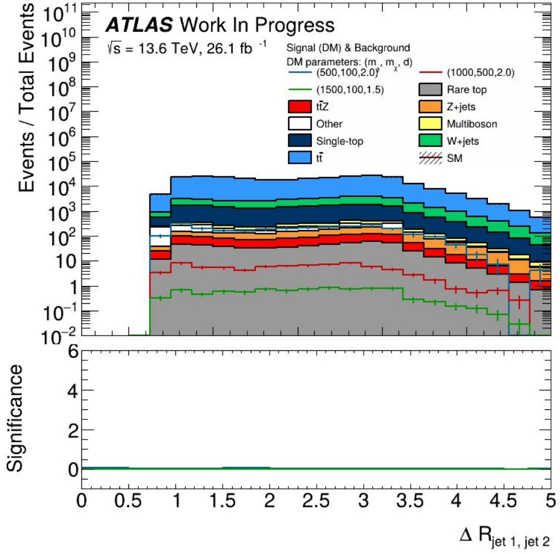




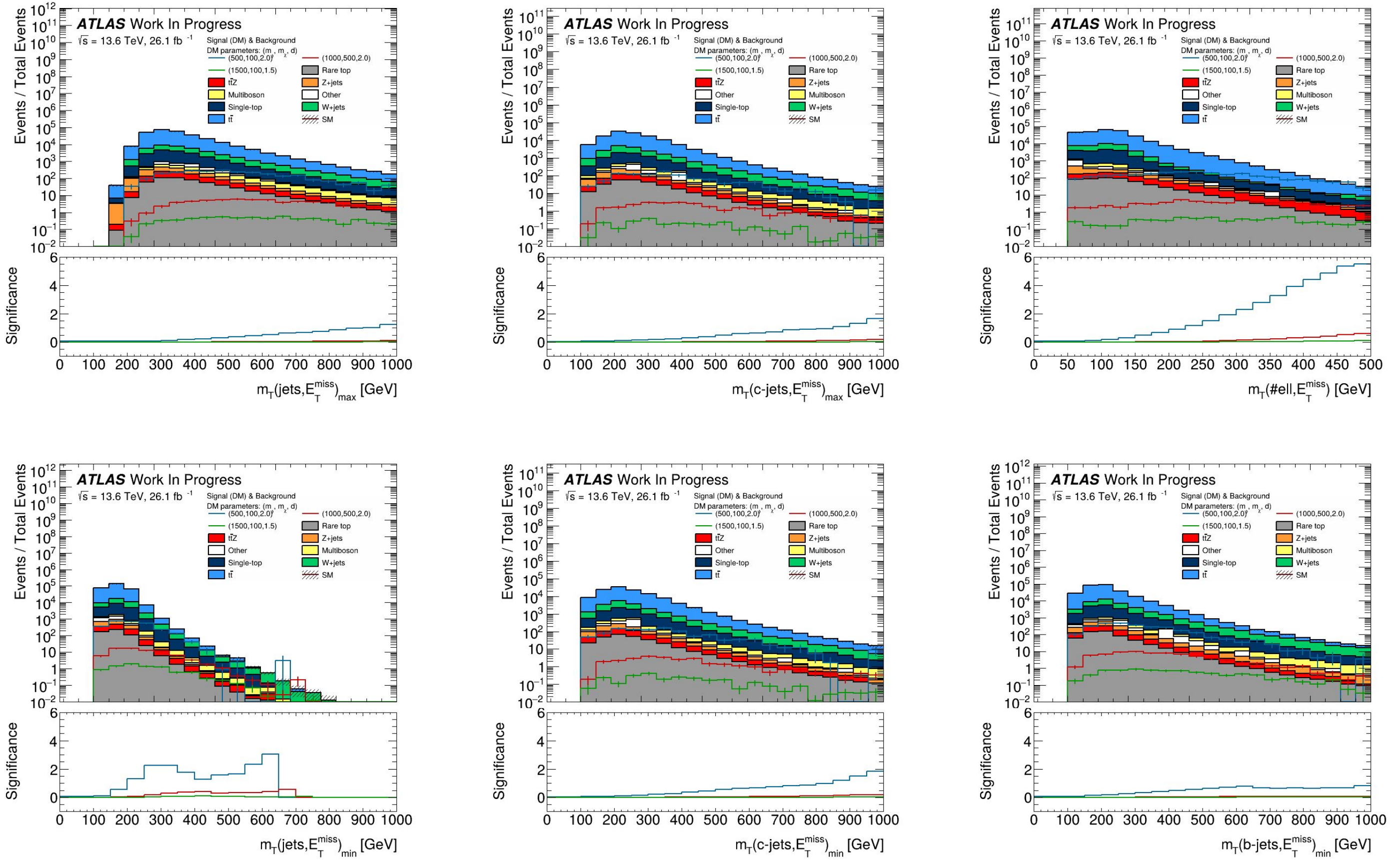
Angular distances



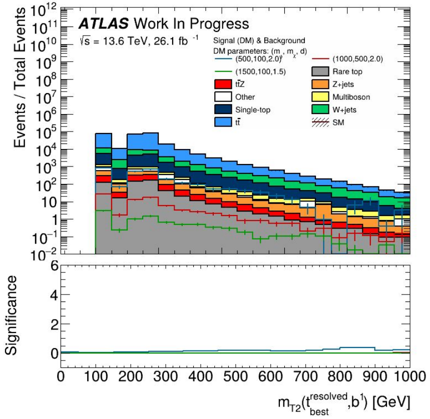


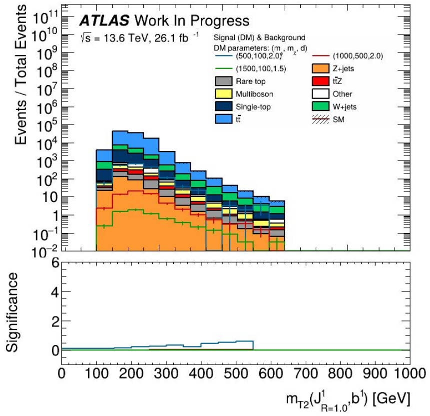


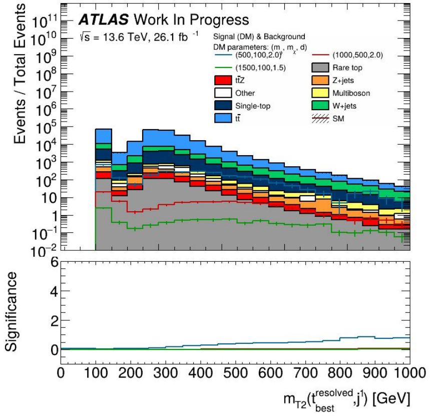
Transverse masses

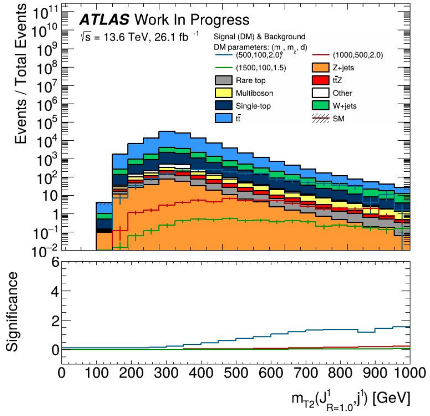


Stransverse masses





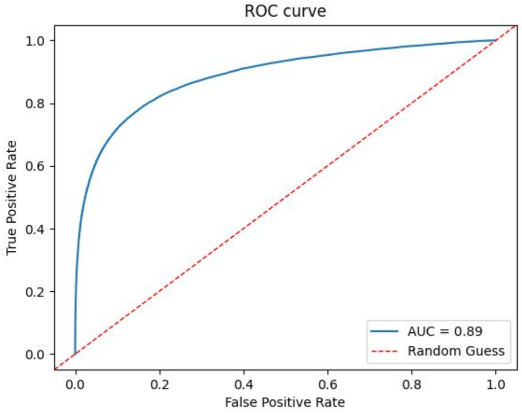


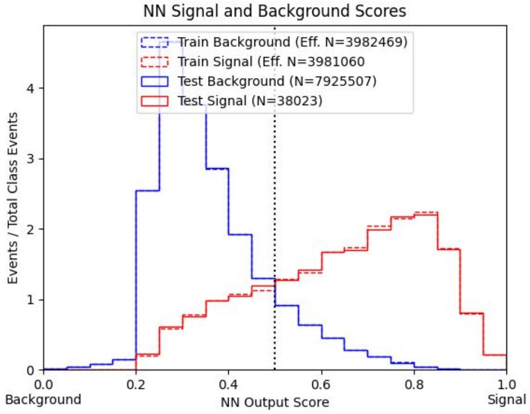


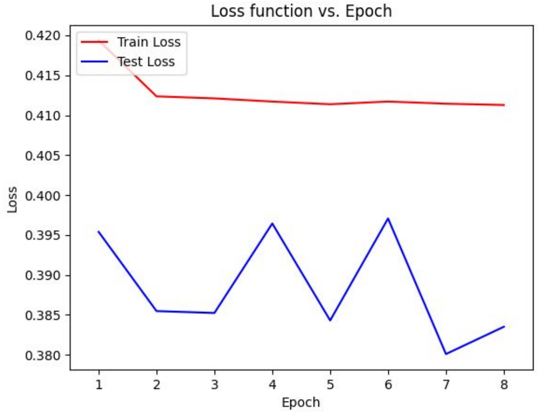
Back-up

Neural Network variable set performances

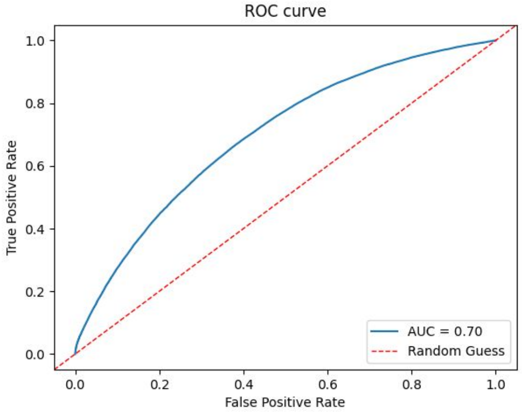
I. Basic kinematics: $E_T^{miss} + p_T^{jets}$

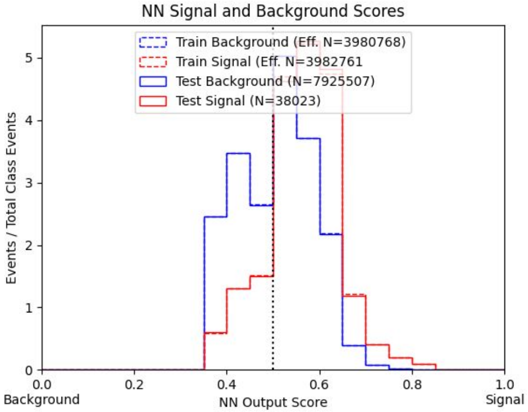


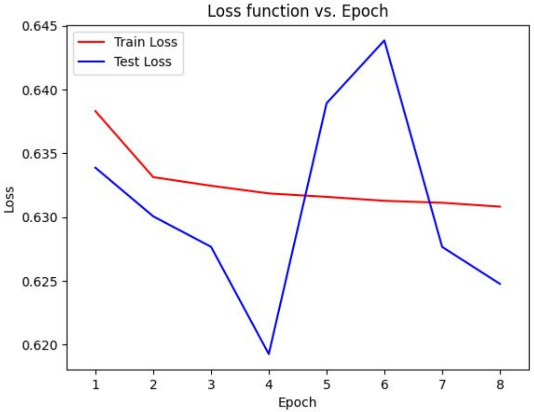




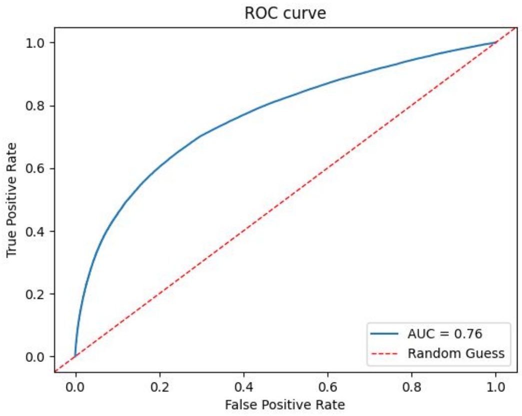
II. Tagged kinematics: $p_T^{b\text{-jet}} + p_T^{c\text{-jet}}$

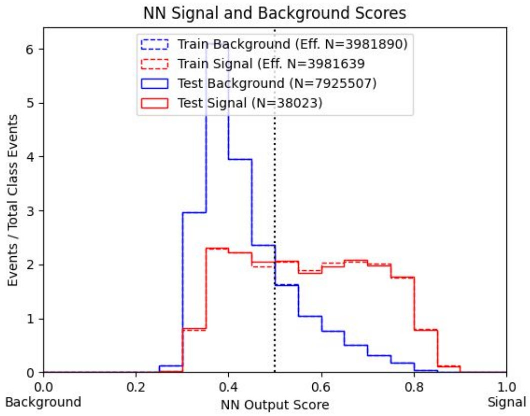


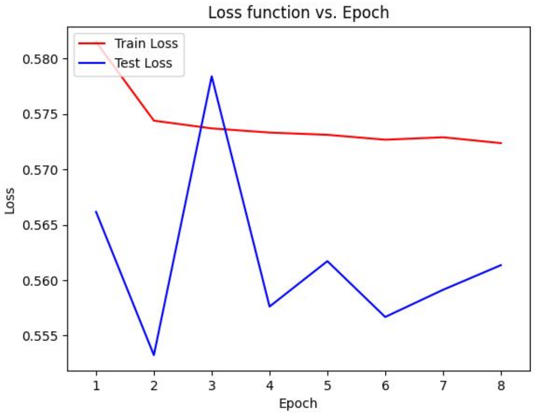




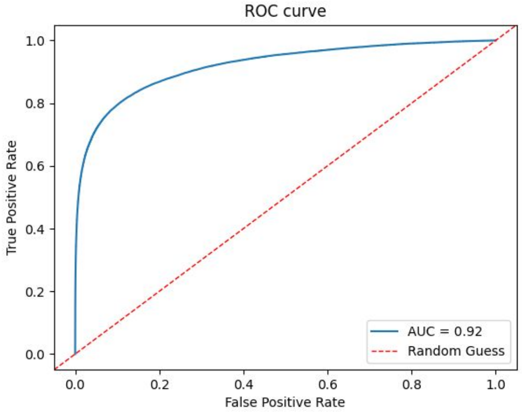
III. Angular distances: ΔR

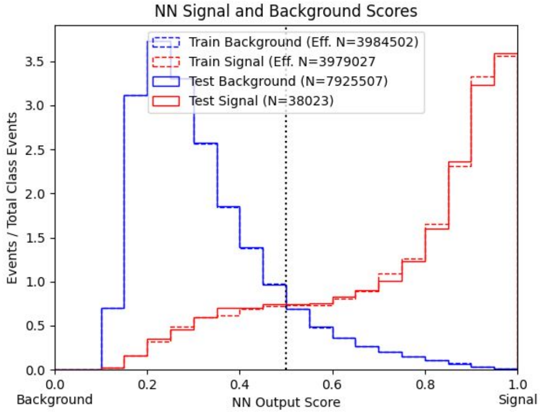


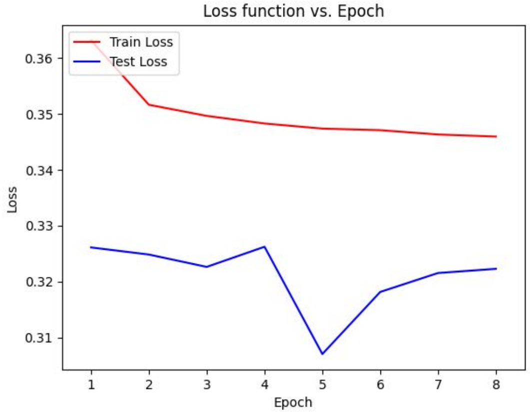




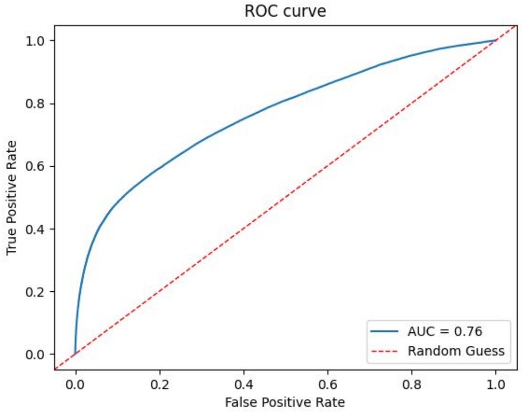
IV. Transverse masses: m_T

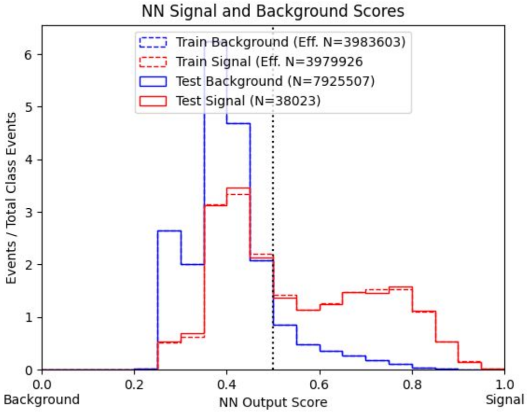


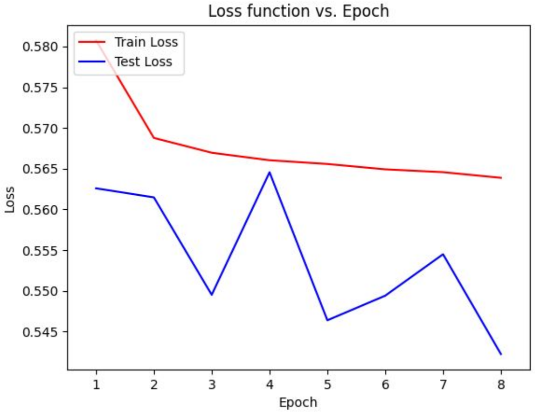




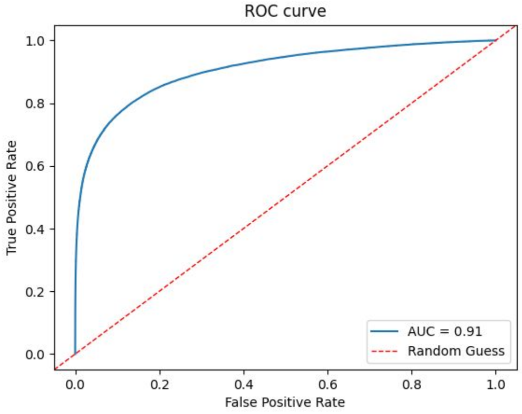
V. Stransverse masses: m_{T2}

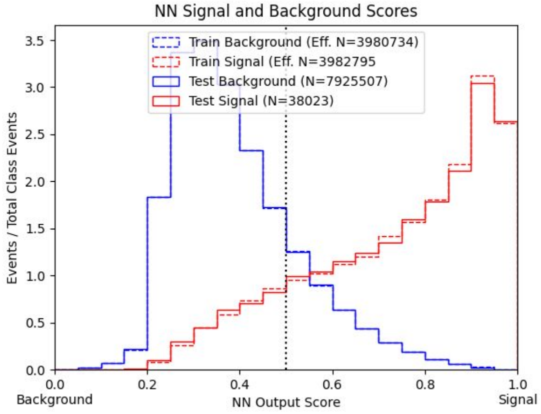


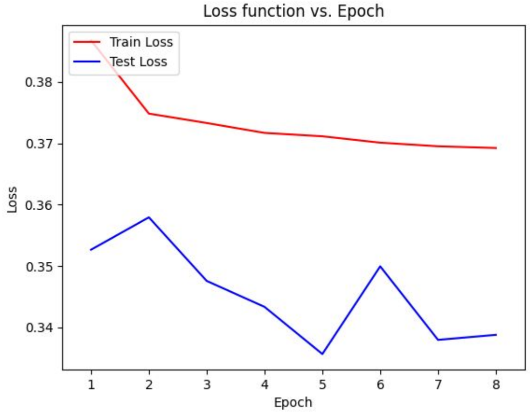




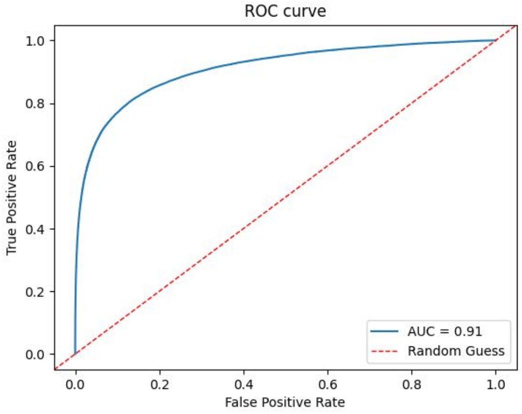
VI. Basic kinematics + Angular distances

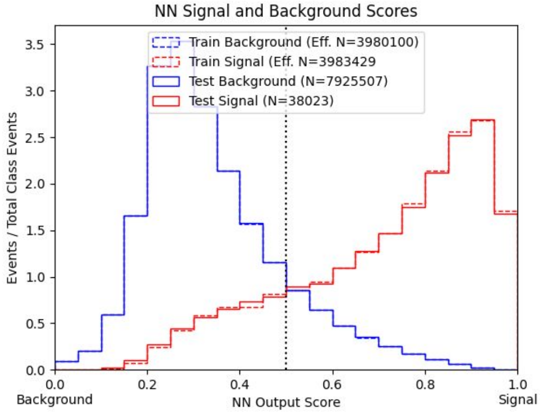


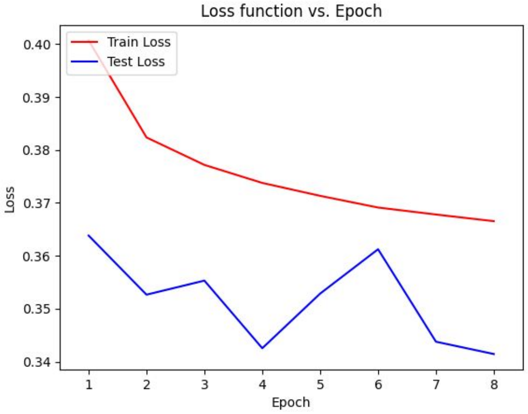




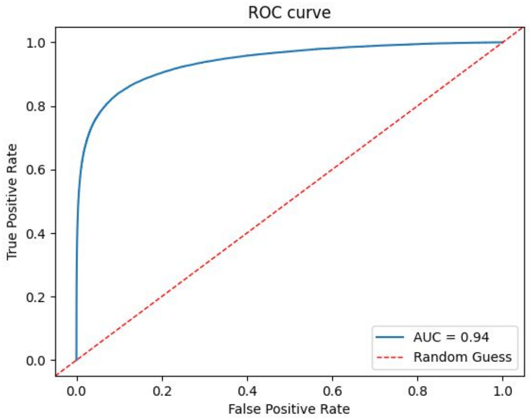
VII. Basic kinematics + Masses

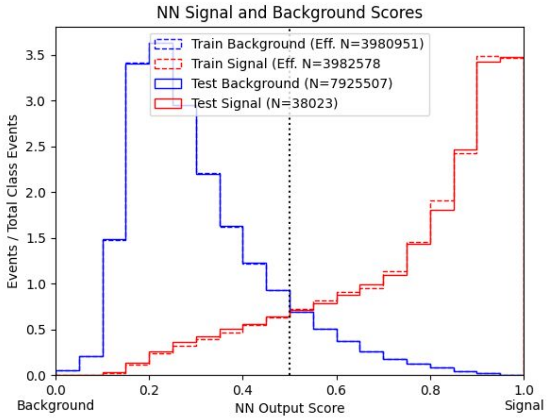


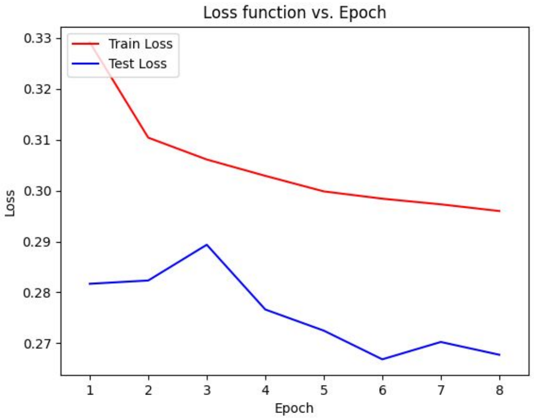




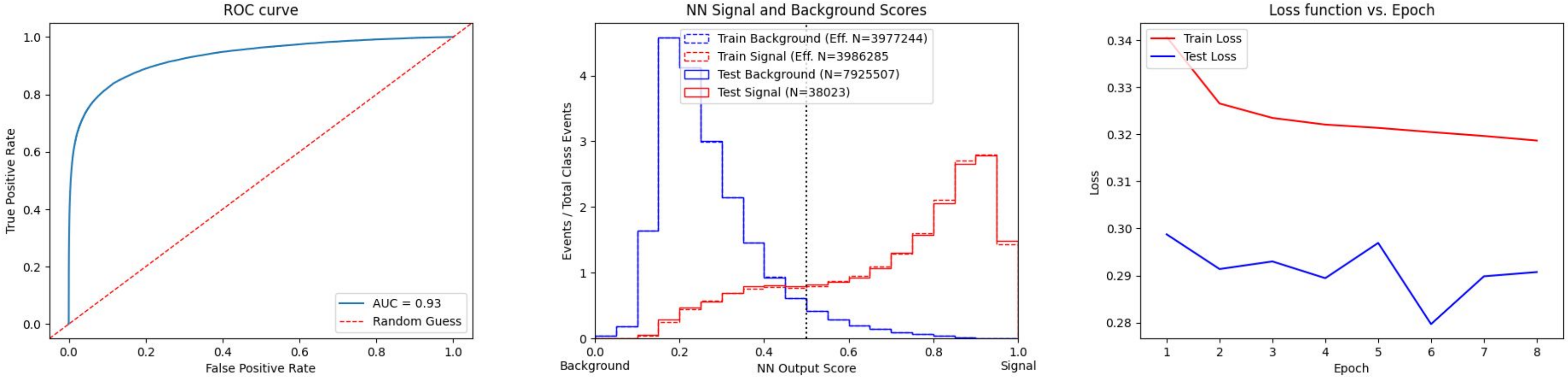
VIII. ALL: Basic & Tagged kin. + Angular + Masses







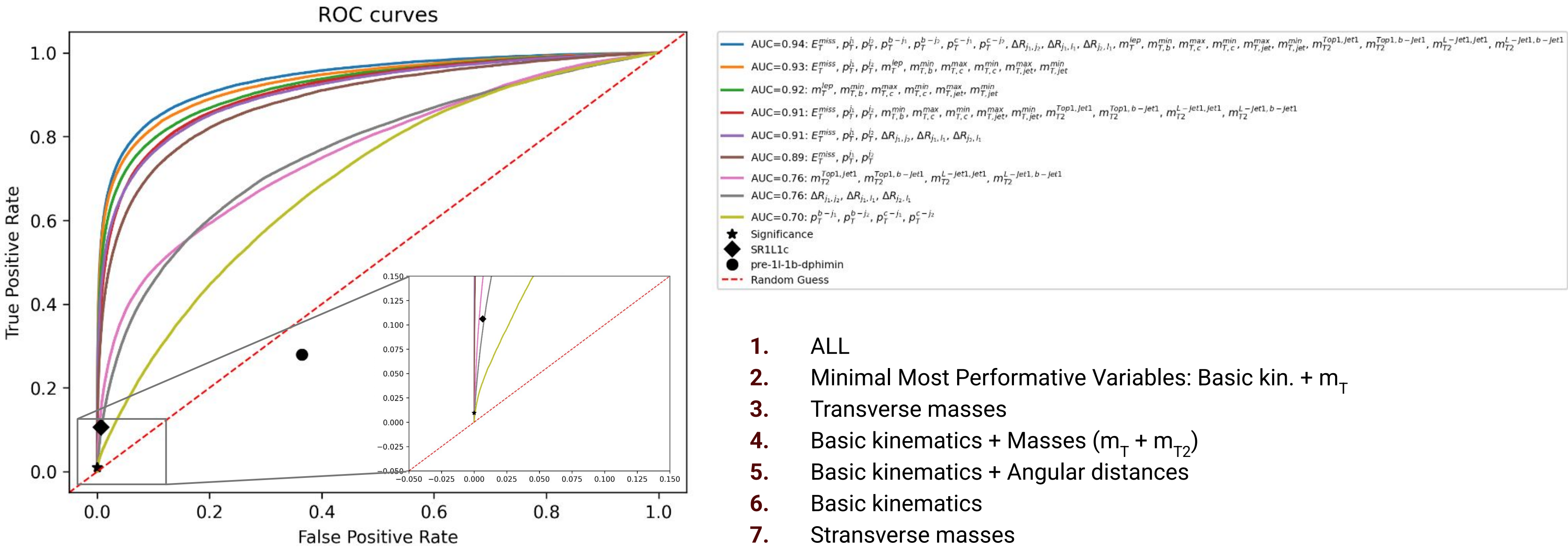
IX. Minimal Most Performative Variables: MET + p_T + m_T



Back-up

Some other slides

“Comparation”: classical cuts vs. DNN



1. ALL
2. Minimal Most Performative Variables: Basic kin. + m_T
3. Transverse masses
4. Basic kinematics + Masses (m_T + m_{T2})
5. Basic kinematics + Angular distances
6. Basic kinematics
7. Stransverse masses
8. Angular distances
9. Tagged kinematics

Reco & *Truth* Monte Carlo simulations

***Truth* simulations:** Events exactly how they would be produced at the centre of the detector, the theoretical events without the effects of the detector (i.e. particles that escape the detector, for example)

***Reco* simulations:** The events of the *truth* simulations are passed through another algorithm that simulates the effects of the detector on the data. This simulations is what we really expect to obtain from our measures.