

Search for compressed mass Higgsino production with soft lepton tracks with the CMS experiment in proton-proton collision data at $\sqrt{s} = 13$ TeV

VON DER FAKULTÄT FÜR PHYSIK DER UNIVERSITÄT HAMBURG ZUR
ERLANGUNG DES AKADEMISCHEN GRADES EINES DOKTORS DER
NATURWISSENSCHAFTEN GENEHMIGTE DISSERTATION

VORGELEGT VON

YUVAL NISSAN

IM AUGUST 2022



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Abstract This is the abstract

Contents

1	Introduction	1
2	Quantum Field Theory and The Standard Model	3
2.1	Quantum Field Theory	3
2.2	The Standard Model of particle physics	3
2.2.1	The particle content	3
3	Supersymmetry	5
3.0.1	Phenomenology of Higgsino production	5
4	Multivariate Statistics	7
4.1	Decision Trees	7
5	Experimental setup	9
5.1	The Large Hadron Collider	9
5.2	The Compact Muon Solenoid experiment	9
5.3	Simulation of events	9
6	Object reconstruction and identification of particles and jets	11
6.1	Track reconstruction	11
6.2	Electrons	11
6.3	Muons	11
6.4	Jets	11
7	Search for compressed Higgsinos with soft lepton tracks	13
7.1	Motivation	13
7.2	Signal models	13
7.3	Previous searches	13
7.4	Search strategy	14
7.4.1	Exclusive track category	14
7.4.2	Dimuon category	14
7.5	Signal signature and base selection	15
7.5.1	Missing Transverse Energy	15
7.5.2	Jets and hardronic activity	15
7.5.3	Base selection	16
7.5.4	Dilepton kinematics	17
7.5.4.1	Lepton η and transverse momentum p_T	18
7.5.4.2	Invariant mass $m_{\ell\ell}$	20
7.5.4.3	Lepton separation ΔR	23
7.5.5	Main drivers of sensitivity	23
7.6	Simulated samples	25

7.6.1	Standard Model simulated samples	25
7.6.2	Signal simulated samples	25
7.7	Object definition and selection	26
7.7.1	Electrons	26
7.7.2	Muons	31
7.7.3	Scale factors	35
7.7.4	Missing transverse energy	40
7.7.5	Jets	40
7.7.6	Tracks and multivariate selection	41
7.7.7	Isolation	45
7.8	Trigger	49
7.9	Event selection	50
7.9.1	Baseline selection	50
7.9.2	Category selection	51
7.9.2.1	Dilepton selection	51
7.9.2.2	Exclusive track selection	51
7.9.3	Boosted decision trees	52
7.9.3.1	Dimuon category	52
7.9.3.2	Exclusive track category	55
7.10	Characterization and estimation of the Standard Model backgrounds	58
7.10.1	Characterization of the Standard Model backgrounds	58
7.10.1.1	Dimuon category	60
7.10.1.2	Exclusive track category	62
7.10.2	Estimation of the Standard Model backgrounds	65
7.10.2.1	Non-isolated jetty background estimation	66
7.10.2.2	Ditau Drell-Yann background estimation	68
7.10.2.3	Exclusive track background estimation	71
7.11	Data control region plots	73
7.12	Signal regions definition and optimization	74
7.13	Systematic uncertainties	76
7.14	Data quality	77
7.15	Results and interpretation	78
8	Summary	79
9	Latex stuff	81
9.1	Some examples	81
9.1.1	Multiline comment	81
9.1.2	Fixme note	81
9.1.3	Tables	81
9.1.4	Cross References	81
9.1.5	Particles	81
9.1.6	Citing	81
9.1.7	Glossary	81
9.1.8	Acronyms	82
9.1.9	fractions	82
	Glossary	85
	Acronyms	87

Chapter 1

Introduction

This is a line in introduction. This is the introduction to the thesis.

Chapter 2

Quantum Field Theory and The Standard Model

2.1 Quantum Field Theory

2.2 The Standard Model of particle physics

2.2.1 The particle content

Chapter 3

Supersymmetry

3.0.1 Phenomenology of Higgsino production

Chapter 4

Multivariate Statistics

4.1 Decision Trees

Chapter 5

Experimental setup

One of the most useful methods to study the subatomic world of particle physics is by using particle accelerators. In such machines, particles are accelerated to very high speeds and energies, and smashed into each other. The resulted particles are then studied by measuring them in a detector. At the time of writing this thesis, the largest currently active accelerator is the Large Hadron Collider (LHC) near Geneva, Switzerland, operated by the European Organization for Nuclear Research (CERN). For the analysis presented in this thesis, data from the Compact Muon Solenoid (CMS) experiment has been analyzed. In this chapter, the LHC is described in 5.1, while the CMS experiment is described in 5.2.

5.1 The Large Hadron Collider

5.2 The Compact Muon Solenoid experiment

5.3 Simulation of events

Chapter 6

Object reconstruction and identification of particles and jets

6.1 Track reconstruction

This is the track reconstruction page.

6.2 Electrons

6.3 Muons

6.4 Jets

Chapter 7

Search for compressed Higgsinos with soft lepton tracks

7.1 Motivation

7.2 Signal models

The signal models considered in this analysis are based on **FiXme Note: fill in signal model stuff.**

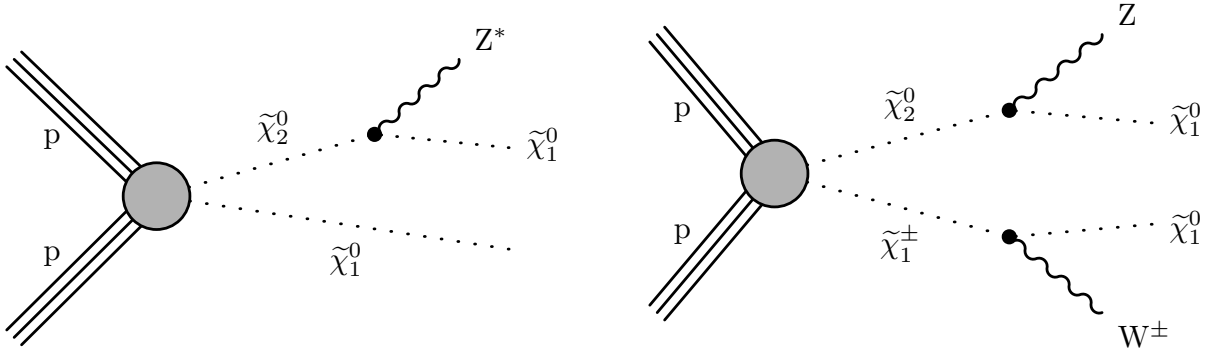


Figure 7.1: Production and decay of electroweakinos in the higgsino simplified model through $\tilde{\chi}_2^0 \tilde{\chi}_1^0$ (left) and $\tilde{\chi}_2^0 \tilde{\chi}_1^\pm$ (right).

FiXme Note: make sure we define the different deltaM somewhere

7.3 Previous searches

7.4 Search strategy

7.4.1 Exclusive track category

7.4.2 Dimuon category

7.5 Signal signature and base selection

To build an effective analysis strategy, the signal kinematics must be studied and exploited. The electroweakino production in question exhibit unique features which can be used in order to discriminate between the signal and the Standard Model (SM) background. It is important to explore these signal distributions in order to define a preselection, or a base cut, that will serve the purpose of retaining as much signal as possible while rejecting as much background as possible. All of the following distributions were plotted by weighting the simulation to Run II luminosity of $\mathcal{L} = 135 \text{ fb}^{-1}$ and requiring at least one jet in the event with $p_T \geq 30 \text{ GeV}$ and $|\eta| < 2.4$. Further selection might apply and will be listed in each section in that case.

7.5.1 Missing Transverse Energy

One property that essentially all Dark Matter (DM) searches have in common is the presence of a DM candidate in the production. The exact identity and properties of said particle (or particles in the case multiple DM candidates) vary, but they do share a lot in common. The DM candidate in our Supersymmetry (SUSY) search is the neutralino, which is a type of DM candidate referred to as a Weakly Interacting Massive Particle (WIMP). A WIMP, broadly speaking, is a new elementary particle which interacts via gravity and any other force (or forces), potentially not part of the SM itself, which is as weak as or weaker than the weak nuclear force, but also non-vanishing in its strength. That essentially means that such candidate is neutral, and therefore not interacting via the electromagnetic force. A neutral particle that interacts neither electromagnetically nor via the strong force (i.e. colorless) will escape detection and will leave traces in the form of a transverse momentum imbalance, which we refer to as E_T^{miss} (Missing Transverse Energy or Missing Transverse Momentum). Our signal contains two DM candidates in the production, which are the Lightest SUSY Particles (LSPs), the neutralinos $\tilde{\chi}_1^0$. We therefore expect the signal to contain considerable magnitude of E_T^{miss} . As described in 7.7.4, we are more interesting in H_T^{miss} , which is highly correlated with E_T^{miss} , due to our definition of lepton isolation and its use in the background estimation methods. Nonetheless, we will look at both E_T^{miss} and H_T^{miss} observables in 7.2.

As we expect, E_T^{miss} and H_T^{miss} are hardly affected by the different choices for Δm , while the higgsino parameter μ affects the distributions above all through its lower production cross section for higher higgsino parameter μ . As discussed at 7.8, the region of interest lie at $H_T^{\text{miss}} \geq 220$ for triggering purposes. Even though this is quite a harsh and inefficient cut, one must look also at the SM background at the regions of $H_T^{\text{miss}} < 220$ and $H_T^{\text{miss}} \geq 220$ to conclude that most of the sensitivity comes from the $H_T^{\text{miss}} \geq 220$ region, since the production of real H_T^{miss} (or E_T^{miss}) results from the production of neutrinos in the event, and these are much less common than Quantum Chromodynamics (QCD) events which swarm the $H_T^{\text{miss}} < 220$ region. Therefore, cutting at $H_T^{\text{miss}} \geq 220$ might be inefficient, but results in high sensitivity.

7.5.2 Jets and hardronic activity

Since the neutralinos $\tilde{\chi}_1^0$ escaping the detector are the contributors to the H_T^{miss} and in doing so the drivers of the sensitivity in high H_T^{miss} region, we want them to be as boosted as possible, i.e., with the highest transverse momentum p_T as possible. A widely used approach is to require an Initial State Radiation (ISR) jet in the event. An ISR jet is formed when one of the incoming protons emit radiation (such as a photon or a gluon) before the interaction. If a jet with high enough p_T is emitted, the rest of the interaction is recoiled against this jet and boosts it in the

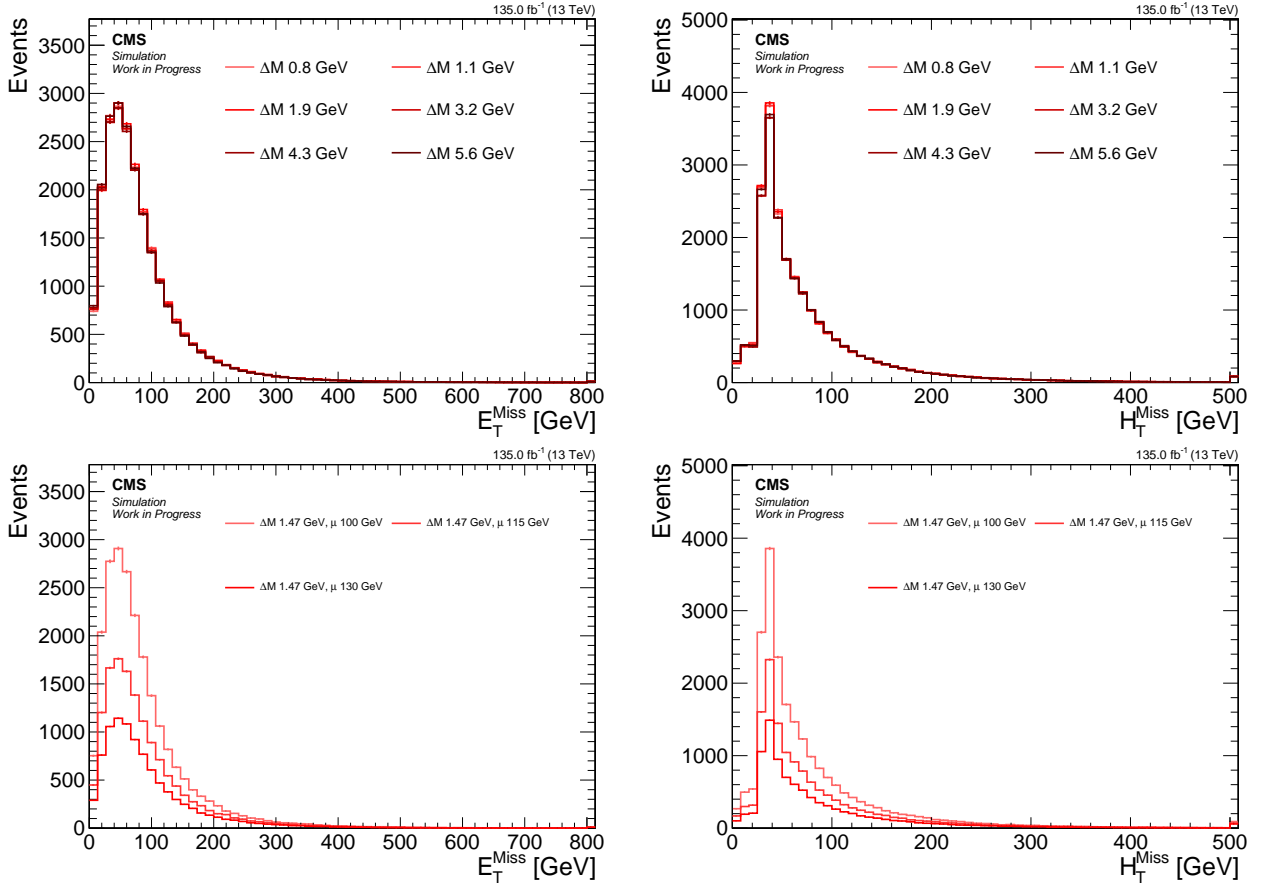


Figure 7.2: Signal distributions for E_T^{miss} (left) and H_T^{miss} (right) comparing various Δm with a fixed higgsino parameter $\mu = 100$ GeV (upper), and comparing various higgsino parameters μ with fixed $\Delta m = 1.47$ GeV (lower).

other direction. This way, the boosted neutralinos $\tilde{\chi}_1^0$ will result in higher H_T^{miss} . As described in 7.7.5, we require the jets to have $p_T \geq 30$ GeV and be located within the tracker acceptance ($|\eta| < 2.4$). We require at least one such jet in the event. The distributions of the number of jets and the leading jet p_T are seen in 7.3.

Our signal signature does not include a b-jet, i.e., a jet resulting from a bottom quark hadronization (either resulting from a top quark or not). We therefore seek to exploit this knowledge by vetoing b-tagged jets in the event. As described in 7.7.5, we are using DEEPCSV flavor tagging discriminant with a medium working point. As can be seen in these distributions, most of the signal lie in the 0 bin, and we will therefore veto any b-tagged jet, which retains most of the signal, but rejects a lot of SM background, such as arising from $t\bar{t}$ events.

Since we are requiring an ISR jet in the event, we expect the E_T^{miss} and the H_T^{miss} to point in the opposite direction of the jet, or at least in an angle close to π . Events with multiple jets in the SM background such as arising from QCD will not exhibit such a feature. In order to reduce QCD background, we require $\min \Delta\phi(H_T^{\text{miss}}, \text{jets}) > 0.4$.

7.5.3 Base selection

We recap the section by summarizing the base selection of our analysis. This base selection, or preselection as we might sometimes call it interchangeably, is applied to all analysis categories.

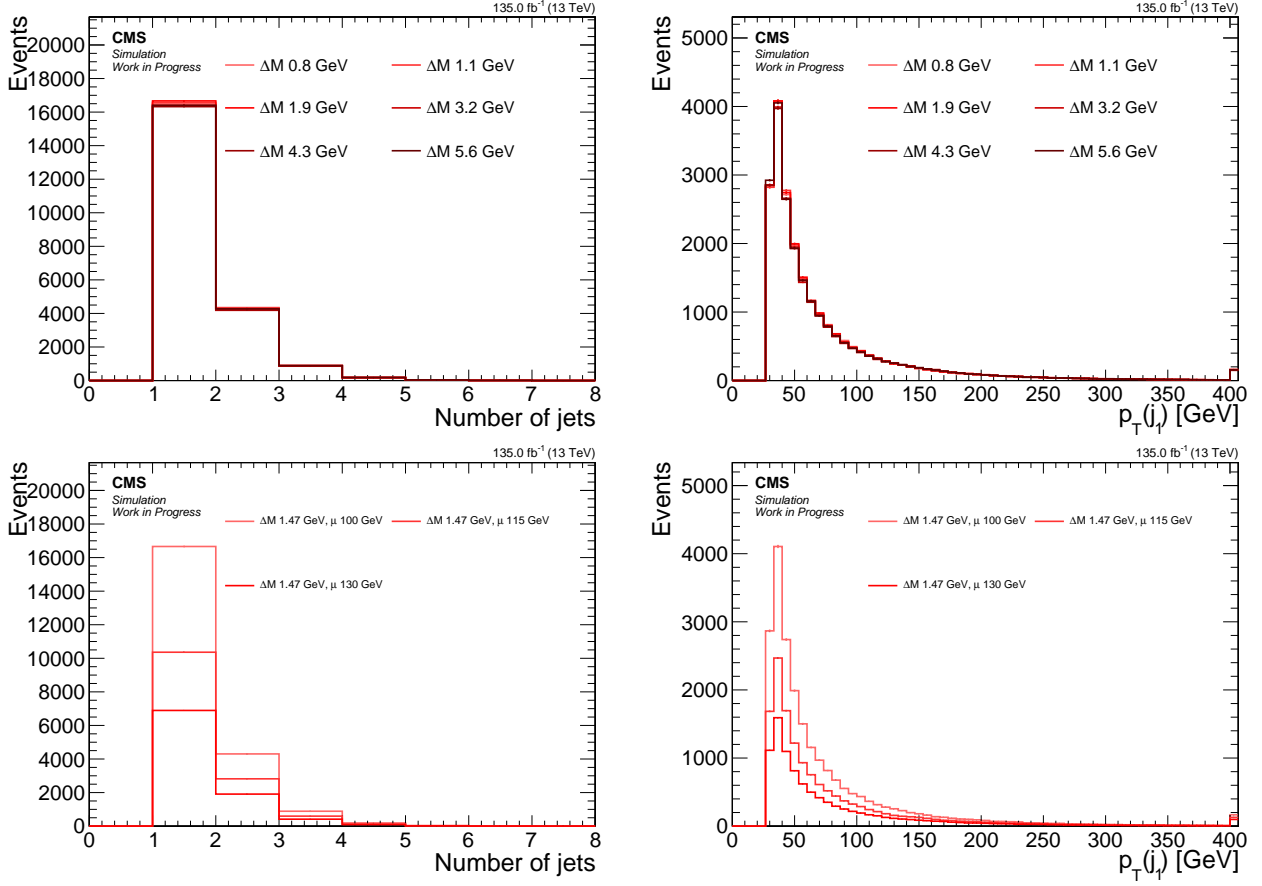


Figure 7.3: Signal distributions for *number of jets* (left) and *leading jet p_T* (right) comparing various Δm with a fixed higgsino parameter $\mu = 100$ GeV (upper), and comparing various higgsino parameters μ with fixed $\Delta m = 1.47$ GeV (lower).

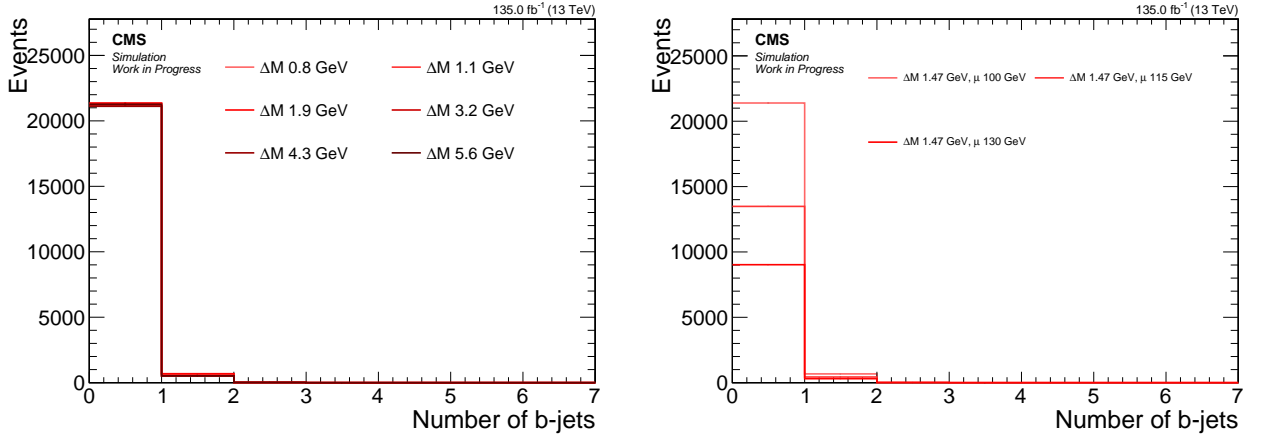


Figure 7.4: Signal distributions for *number of b -tagged jets* comparing various Δm with a fixed higgsino parameter $\mu = 100$ GeV (left), and comparing various higgsino parameters μ with fixed $\Delta m = 1.47$ GeV (right).

7.5.4 Dilepton kinematics

Thus far we have looked at kinematic distributions ignoring the leptons in event. However, the most distinct features of the signal lie in the dilepton system. To fully understand the unique phase space of the dilepton system, we first look at generator level distributions and then look at what effects does reconstruction have on those observables. In addition, since the most sensitive category is the dimuon category, due to its much lower threshold on the

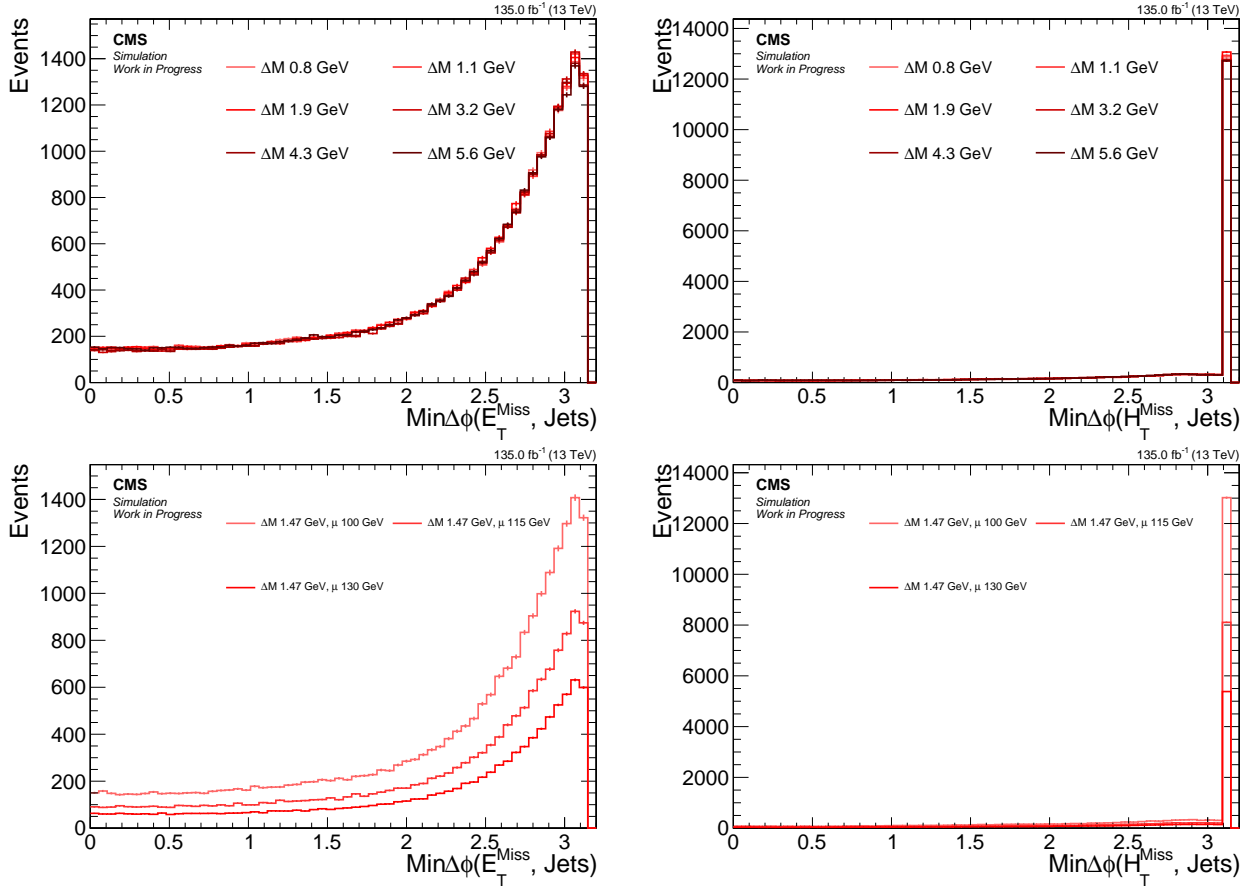


Figure 7.5: Signal distributions for $\min \Delta\phi(E_T^{\text{miss}}, \text{jets})$ (left) and $\min \Delta\phi(H_T^{\text{miss}}, \text{jets})$ (right) comparing various Δm with a fixed higgsino parameter $\mu = 100$ GeV (upper), and comparing various higgsino parameters μ with fixed $\Delta m = 1.47$ GeV (lower).

Table 7.1: Base selection applied to all analysis categories

Variable	Value
H_T^{miss} [GeV]	≥ 220
$N_{\text{jets}} (p_T \geq 30 \text{ GeV and } \eta < 2.4)$	≥ 1
$N_{\text{b-jets}} (p_T \geq 30 \text{ GeV and } \eta < 2.4)$	0
$\min \Delta\phi(H_T^{\text{miss}}, \text{jets})$	> 0.4

transverse momentum, these events are shown in the following sections, leaving out events with two electrons.

Since the kinematics change dramatically as a function of Δm , but change almost only in overall normalization due to the production cross section as a function of the higgsino parameter μ , in the following sections, we set the higgsino parameter to $\mu = 100$ GeV and vary the Δm .

7.5.4.1 Lepton η and transverse momentum p_T

The transverse momentum p_T distribution of the muons and our access to a subset of its full range have a dramatic effect on the signal acceptance and the sensitivity. The muon reconstruction process and details are discussed in 6.3. The selection we apply to the muons in this analysis is described in 7.7.2 and will be referred to here as *analysis selection*. In this section we would like to consider specifically the importance of the p_T on the signal and its dilepton kinematic distributions.

We begin by having a look at the generator level distribution of p_T , or in other words, the so-called *truth* distributions, which do not exhibit any detector or reconstruction features, and compare then to the reconstructed distribution, as seen in 7.6.

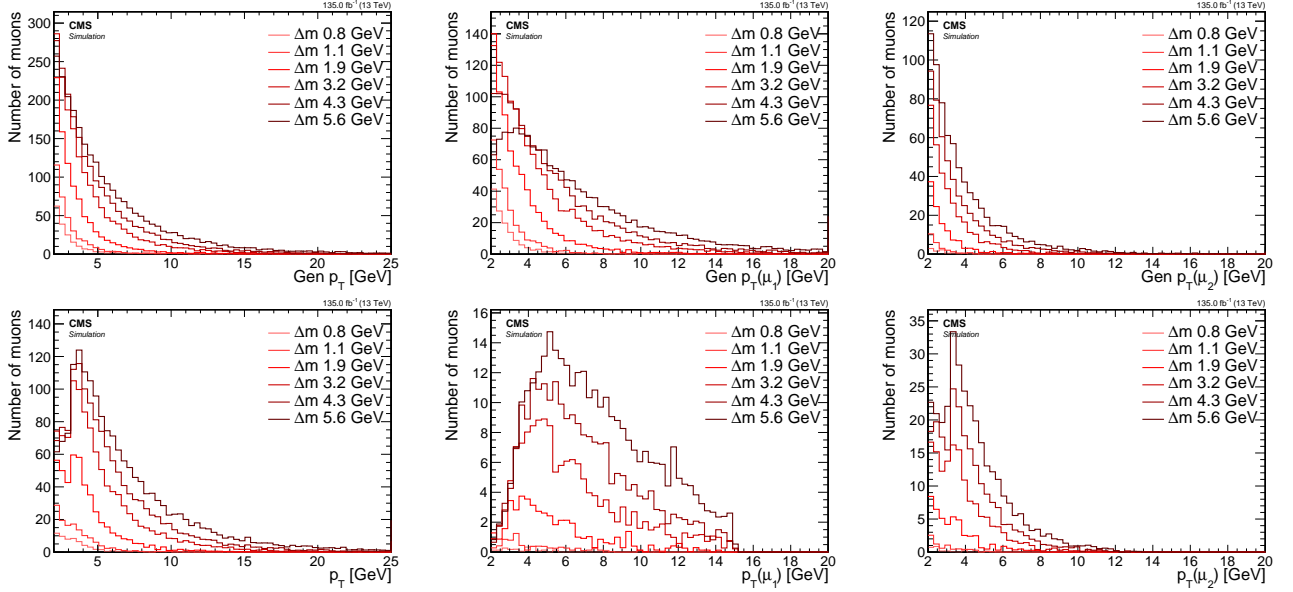


Figure 7.6: Signal p_T distributions for inclusive (left), leading muon μ_1 (middle), subleading muon μ_2 (right) at generator level (top) and reconstruction level passing analysis selection (bottom).

By comparing the generator level to the reconstruction level of the inclusive p_T distribution, we see that a reshaping occurs at around 3 GeV. A significant portion of the generated muons with $p_T < 3$ GeV are being lost in reconstruction. The reconstructed subleading p_T distribution has a camel shape whereby the efficiency drops below p_T of 3 GeV and only partially regained at $p_T < 3$ GeV. This is a detector effect and can be seen more clearly when we split the p_T distribution into a barrel ($|\eta| < 1.2$) and endcaps ($|\eta| \geq 1.2$) portions, as is shown in 7.7.

The picture becomes much clearer in regards to the reconstruction efficiency as a function of p_T . When comparing the generator level distribution of the barrel muons on the top left with its constructed counterpart on the bottom left, we see that the barrel is almost completely unable to reconstruct muons with $p_T < 3$ GeV, while the endcaps, shown on the left, are able to do so. As we will see in the $m_{\ell\ell}$ and ΔR upcoming sections (7.5.4.2 and 7.5.4.3), since those distribution have an important relationship, that has consequences in regards to reshaping kinematic distributions, as well as signal acceptance in general. Since the low region of $2 \leq p_T \leq 3.5$ GeV is crucial in giving us access to low Δm signal points, it is achieved, as can be seen here, mainly with the help of the muon chamber endcaps.

Since the barrel and endcaps are separated by different regions of η , $|\eta| < 1.2$ for barrel and $|\eta| \geq 1.2$ for endcaps, it is worthwhile taking a look at the η distributions of the muons as well. Those can be seen at 7.8.

Our muons analysis selection picks only muons in the tracker range of $|\eta| < 2.4$, which is the reason why the reconstruction plots on the bottom do not have muons with $|\eta| > 2.4$. We see that the main effect going from the inclusive $|\eta|$ in generator level to the reconstructed counterpart, is the flattening of the distribution due to the loss of muons with $|\eta| < 1.2$ in the barrel for muons with $p_T < 3$ GeV.

With the understanding of the reconstruction effects on the p_T and η distributions of the muons, we are equipped to look further into other kinematic variables of the dilepton system.

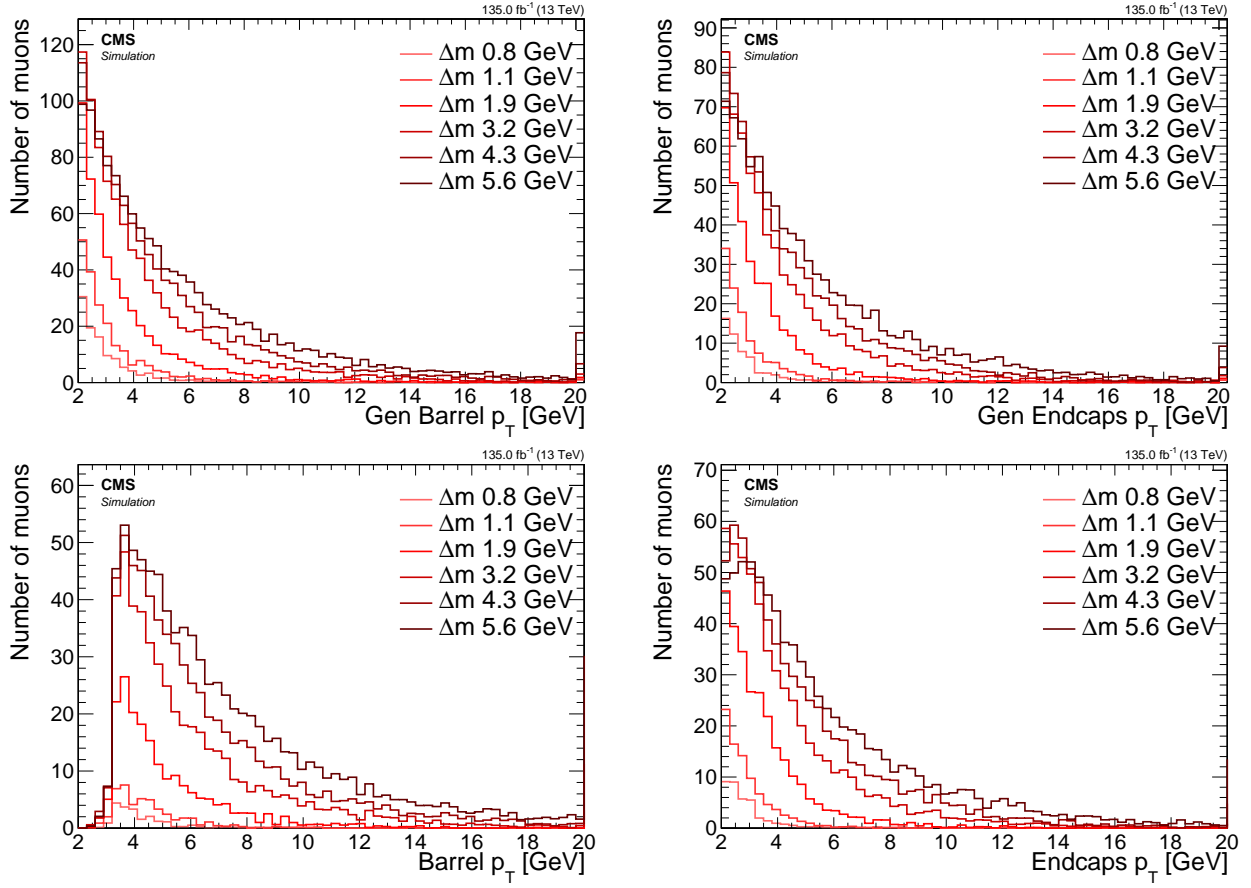


Figure 7.7: Signal inclusive p_T distributions for barrel $|\eta| < 1.2$ (left) and endcaps $|\eta| \geq 1.2$ (right) at generator level (top) and reconstruction level passing analysis selection (bottom).

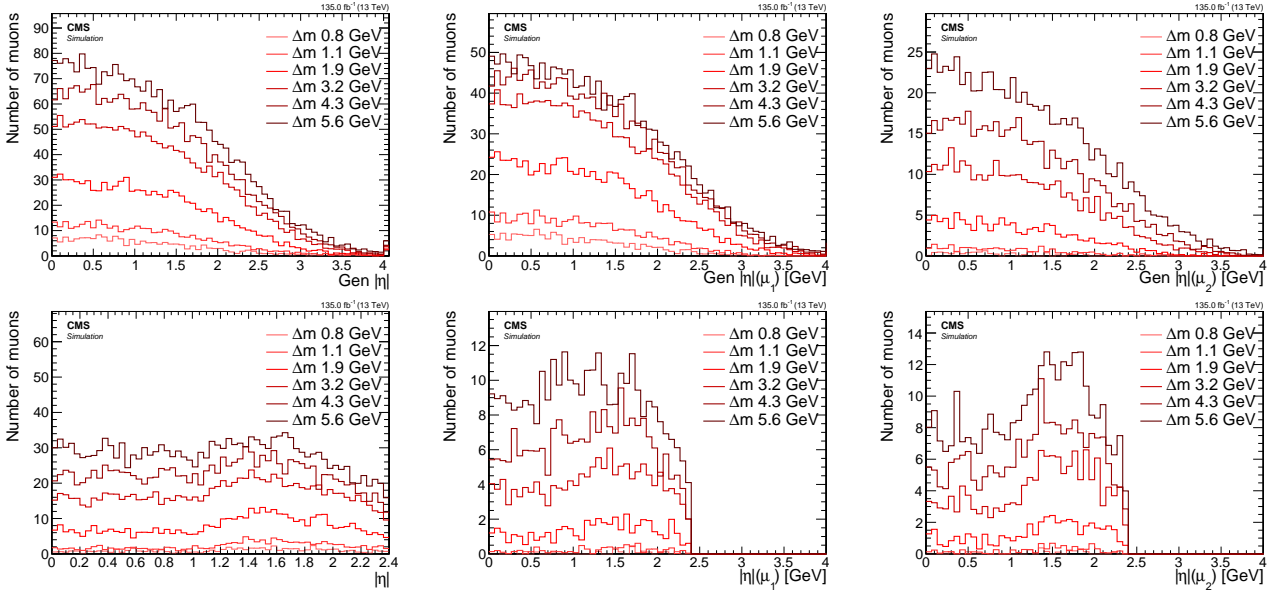


Figure 7.8: Signal $|\eta|$ distributions for inclusive (left), leading muon μ_1 (middle), subleading muon μ_2 (right) at generator level (top) and reconstruction level passing analysis selection (bottom).

7.5.4.2 Invariant mass $m_{\ell\ell}$

The invariant mass of the two leptons that result from the decay of the $\tilde{\chi}_2^0$ has a unique shape due to the limited allowed phase space of the leptons as part of the 3-body decay. Since the $\tilde{\chi}_2^0$

decays into $\tilde{\chi}_1^0$ and $\ell^+\ell^-$ through a Z^* , the allowed phase space of the dilepton pair is restricted to the mass difference between $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$, i.e., Δm . We therefore expect the $m_{\ell\ell}$ distribution to have an edge at Δm . Distributions of the generator level invariant mass are seen in 7.9.

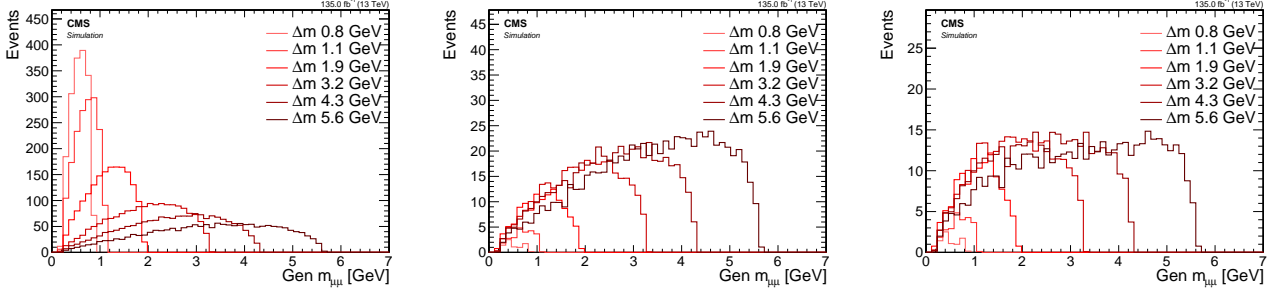


Figure 7.9: Signal generator level $m_{\ell\ell}$ distributions with no cuts (left), with $p_T(\mu_i) > 2 \text{ GeV}$, $i = 1, 2$ (middle) and with SOS orthogonality condition $p_T(\mu_i) > 2 \text{ GeV}$, $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R \leq 0.3$ (right).

We see the inclusive distribution of the invariant mass of the muons $m_{\mu\mu}$ on the left. For each signal point, the edge of the $m_{\mu\mu}$ distribution is right at the corresponding Δm . However, once we cut on the muons p_T and require $p_T \geq 2 \text{ GeV}$, the shape shifts, and the efficiency in the lower Δm drops dramatically, as can be seen in the middle plot. Lastly, the effect of orthogonalizing phase space to the SOS analysis can be seen on the right most plot. The effect is strongest in high Δm and quite subtle in low Δm .

To understand the reshaping that happens to the $m_{\mu\mu}$ shape, we look at the relationship between the p_T of the muons (leading muon denoted μ_1 while subleading muon is denoted μ_2) and the invariant mass. We pick one signal with low Δm of 1.13 GeV and one with high Δm of 5.63 GeV. Those distributions are seen in 7.10.

We have established earlier that the invariant mass distribution has an edge at the Δm and one can read the value of Δm from these plots. Another interesting feature to notice in these plots is that there is also a lower edge in the Δm distribution at around $\sim 0.2 \text{ GeV}$ and that is of course due to each muon having a mass of around $\sim 0.1 \text{ GeV}$. We can clearly see now that by cutting on both muons at $p_T \geq 2 \text{ GeV}$, we are losing a lot of the signal. This effect, in fact, becomes quite substantial for the low $\Delta m = 1.13 \text{ GeV}$ (top row). We quantify this effect by making a cutflow table where each row represents a cut, and its efficiency is calculated by dividing the number of events passing the cut by the number of events in the line before it. The first line is our baseline of all dimuon events with at least one jet with $p_T \geq 30 \text{ GeV}$ and $|\eta| < 2.4$ and has an efficiency of 1 by definition. The event number is weighted to Run II luminosity of $\mathcal{L} = 135 \text{ fb}^{-1}$. The table is seen at 7.5.4.2.

Table 7.2: Generator level efficiency on muons selections

Cut	Number of events		Efficiency	
	$\Delta m = 1.13 \text{ GeV}$	$\Delta m = 5.63 \text{ GeV}$	$\Delta m = 1.13 \text{ GeV}$	$\Delta m = 5.63 \text{ GeV}$
Baseline	1710.7	1743.9	1	1
$p_T \geq 2 \text{ GeV}$	24.7	724.9	0.015	0.41
SOS orthogonality	24.7	490.6	1	0.68

We observe that for the low Δm of 1.13 GeV the cut of $p_T \geq 2 \text{ GeV}$ is really hurting the acceptance of the signal with only 1.5% of signal remaining. In contrast, the orthogonality condition of requiring $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R(\ell\ell) \leq 0.3$ is not effecting it any further. The picture is different for the high Δm of 5.63 GeV where the p_T cut is cutting away more than half of the signal and the SOS orthogonality an additional two thirds.

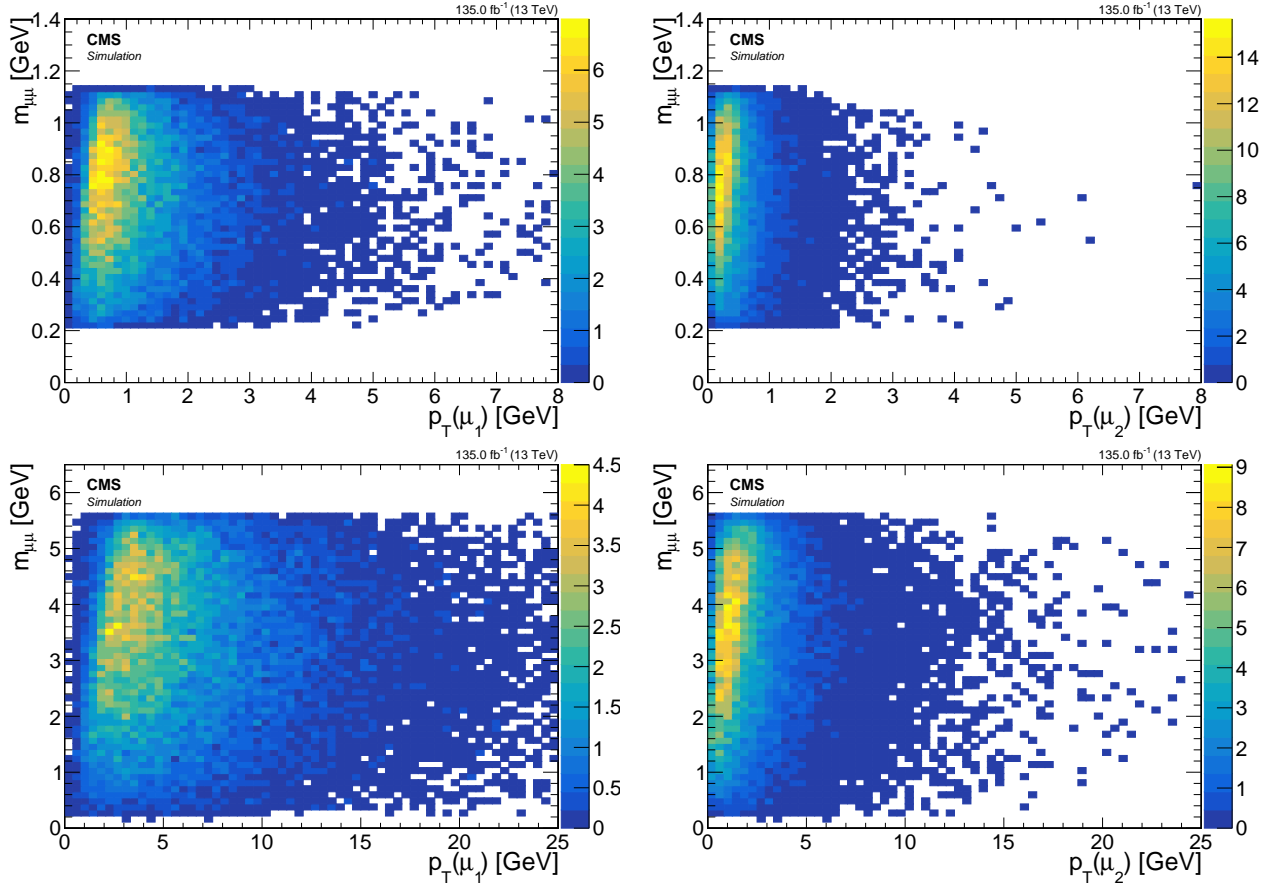


Figure 7.10: Signal $m_{\mu\mu}$ vs. p_T for leading lepton μ_1 (left) and subleading lepton μ_2 (right) for $\Delta m = 1.13$ GeV (top) and $\Delta m = 5.63$ GeV (bottom).

Since we have just established that due to the relationship between p_T and $m_{\ell\ell}$, the p_T distribution directly effects $m_{\ell\ell}$, we should also have a look at how the reconstruction discussed at 7.5.4.1 might effect the $m_{\mu\mu}$ distribution. The distributions of reconstructed $m_{\mu\mu}$ can be seen in 7.11.

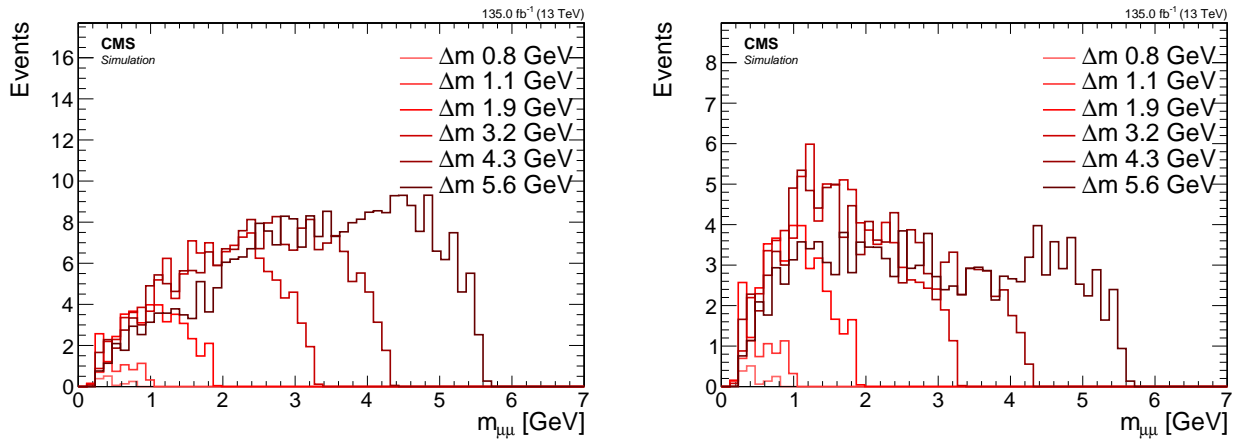


Figure 7.11: Signal reconstructed $m_{\mu\mu}$ with basic analysis selection (left) and additional SOS orthogonality condition (right).

It's interesting to compare these distributions to the two right ones in the generator level version at 7.9. We see that not only less events survive the reconstruction, but also that some Δm model points peak between 1 GeV to 2 GeV with SOS orthogonality condition applied.

7.5.4.3 Lepton separation ΔR

The lepton separation is defined by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ where η is the pseudorapidity and ϕ is the azimuthal angle measured in radians. ΔR plays a major role in this analysis since the leptons tend to be produced in proximity to each other and thus to defy standard definitions of isolation. Special care is taken to ensure that the collimated nature of the leptons can still be used to distinguish the otherwise isolated leptons in the signal from the non-isolated leptons in the SM background. An additional point of interest is with respect to previous SOS analysis [1] that had a requirement of $\Delta R(\ell\ell) > 0.3$ which we attempt to revert here for orthogonality purposes.

In similar fashion to the invariant mass in 7.5.4.2, we look at distributions of ΔR for various choices of Δm with different cuts applied, and observe their effect.

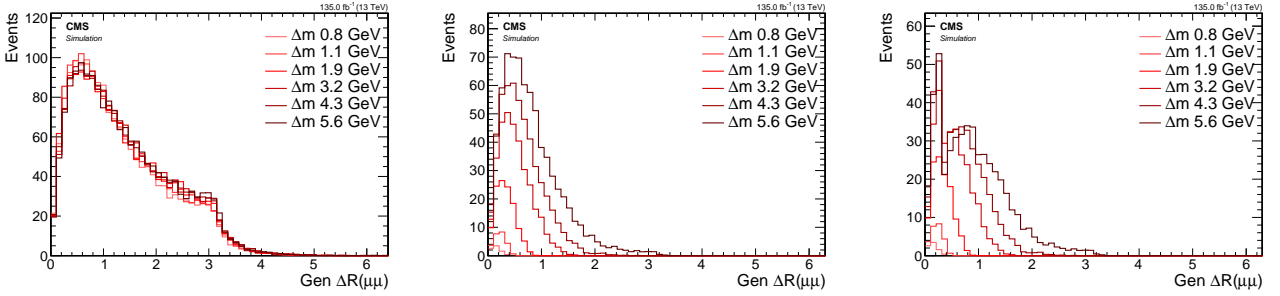


Figure 7.12: Signal generator level ΔR distributions with no cuts (left), with $p_T(\mu_i) > 2$ GeV, $i = 1, 2$ (middle) and with SOS orthogonality condition $p_T(\mu_i) > 2$ GeV, $p_T(\mu_2) \leq 3.5$ GeV or $\Delta R \leq 0.3$ (right).

As can be seen on the left plot, roughly the same amount of events are produced for all Δm model points, but when applying a cut of $p_T(\mu) > 2$ GeV, a hierarchy of Δm points forms, with less events as Δm becomes smaller (middle plot). The spike on the right plot is due to the SOS orthogonality condition which requires $\Delta R(\ell\ell) \leq 0.3$ as one of two conditions in an or statement. To understand the shaping and hierarchy formation due to the p_T cut, we repeat the trick from the $m_{\mu\mu}$ in 7.5.4.2 and plot the p_T of the muons vs. $\Delta R(\ell\ell)$. Those distributions are seen in 7.13.

Now the hierarchy can be understood by observing that for $\Delta m = 1.13$ GeV, cutting on $p_T(\mu_2) \geq 2$ GeV will limit the range of $\Delta R(\mu\mu)$ to less than 0.4 while leaving quite a large range exceeding 3 for the $\Delta m = 5.63$ GeV model point.

We conclude therefore, that even before taking into consideration reconstruction efficiency of the leptons, in order to gain access and sensitivity to the low Δm model points, we must be able to probe low $\Delta R(\ell\ell)$ values, potentially with values less than 0.3. In the next sections we will need to study reconstructed leptons and define isolation criteria that will enable us to retain signal points with close lepton pairs. That is further explored in 7.7.7.

As we have seen for $m_{\mu\mu}$ in 7.5.4.2, reconstruction has an effect on both the shape and overall count of events. We investigate those effects on the $\Delta R(\mu\mu)$ distributions in 7.14.

When we compare the reconstructed $\Delta R(\mu\mu)$ distributions 7.14 to the generator level ones at 7.12 we see that the main effect of the reconstruction on the $\Delta R(\mu\mu)$ is the overall normalization due to reconstruction efficiency.

7.5.5 Main drivers of sensitivity

We attempt to draw conclusions from this signal distribution studies in regards to the main drivers to the sensitivity of different model points of this analysis, as well as of future analysis

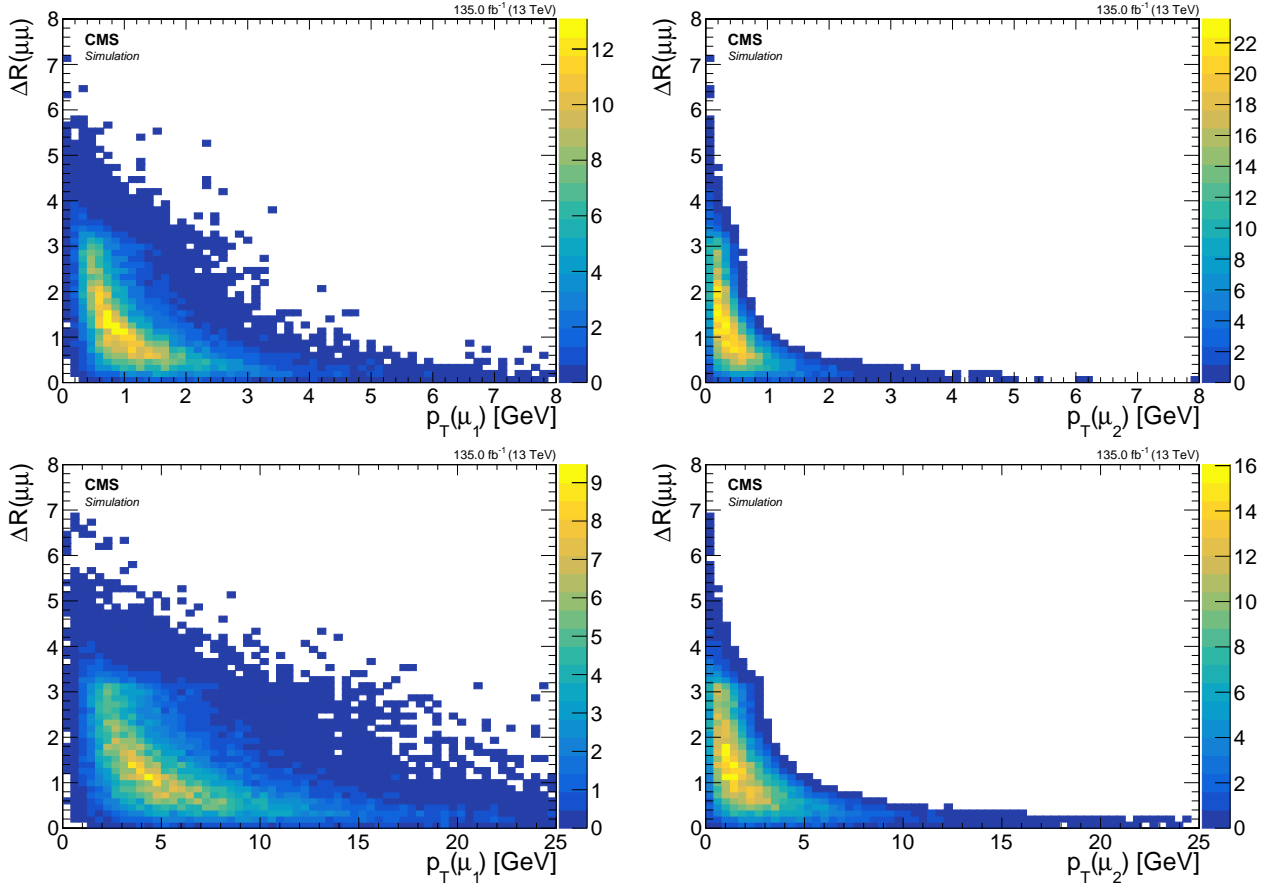


Figure 7.13: Signal $\Delta R(\mu\mu)$ vs. p_T for leading lepton μ_1 (left) and subleading lepton μ_2 (right) for $\Delta m = 1.13$ GeV (top) and $\Delta m = 5.63$ GeV (bottom).

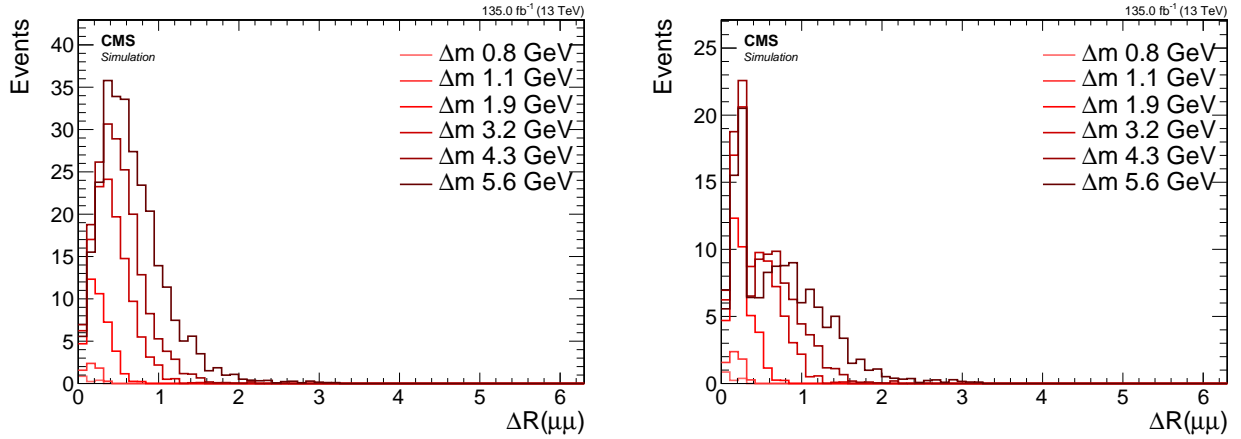


Figure 7.14: Signal reconstructed $\Delta R(\mu\mu)$ with basic analysis selection (left) and additional SOS orthogonality condition (right).

that might try to expend on this one. In this section we have not looked at SM background at all. Therefore, it is hard to conclude what effects changing the cuts to E_T^{miss} or other event level observables might have. However, one thing is very clear from examining the dilepton kinematics, and that is that to gain access to low Δm model points, one must lower the threshold on p_T in the selection of the decay leptons and make sure to be able to also probe low ΔR . Another driver of the sensitivity at all Δm model points is the luminosity, since the production cross section drops as a function of the higgsino mass parameter μ .

We will explore in the next sections how we are able to lower the threshold on the muons

transverse momentum as well as dealing with collimated leptons that might pose a challenge in regards to the isolation criterion.

7.6 Simulated samples

7.6.1 Standard Model simulated samples

7.6.2 Signal simulated samples

7.7 Object definition and selection

We have seen in 6 how objects are being reconstructed and identified in our detector. We have also studied the signal signature in 7.5. In this section we devise an object selection in order to obtain as pure as possible sample of objects in regards to our target leptons, while retaining as much signal as possible. As we have seen in 7.4, we are targeting the opposite-charged same-flavor leptons $\ell^+\ell^-$ that result from the $\tilde{\chi}_2^0$ that decays into a $\tilde{\chi}_1^0$ via a Z^* , i.e., $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$. In the following section, we choose to present two choices of Δm^0 , namely, $\Delta m^0 = 1.92 \text{ GeV}$ and $\Delta m^0 = 5.63 \text{ GeV}$, i.e., a relatively high Δm^0 , and a low one, but not too low as to still be able to have enough electrons surviving the initial reconstruction p_T threshold of 5 GeV. We also fix the higgsino parameter on $\mu = 100 \text{ GeV}$.

As was the case in 7.5, the base selection for the following section is requiring at least one jet in the event with $p_T \geq 30 \text{ GeV}$ and $|\eta| < 2.4$. No other selections otherwise. However, unlike 7.5, we do not weight our objects to any luminosity, as we are interested in the proportion between object types. We differentiate between two types of leptons, ones that originate from our targeted decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$, which will be shown in blue, and those that do not, which we refer to as *other*, and are shown in yellow. Leptons that are marked as resulting from the $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$ decay, which we will refer to as *signal leptons*, are done so by matching a reconstructed lepton to a generator level lepton, which has been checked to have the $\tilde{\chi}_2^0$ as its parent. Lepton marked as *other*, either has been misreconstructed, misidentified or is a result of hadronisation process in a jet (such as the ISR jet). Our goal here is to select as many blue leptons as possible, while rejecting as many yellow ones as possible. In the following sections, we will refer to *efficiency* as the proportion between the signal leptons passing a selection, divided by the initial number of signal leptons, and to *purity* as the proportion between signal leptons (blue) and the sum of the signal leptons and *other* leptons (yellow). So to rephrase our goal, we are interested in a selection that results in high-efficiency and high-purity. These two quantities can sometimes compete with each other and we have to make compromises.

7.7.1 Electrons

The electrons have an initial reconstruction p_T threshold of 5 GeV. The initial working point choice for reconstructed electron is loose (see 6). The first distribution we look at in regards to the electrons is their spatial separation from the leading jet in the event, $\Delta R(j_1, e)$. Those can be seen in 7.15.

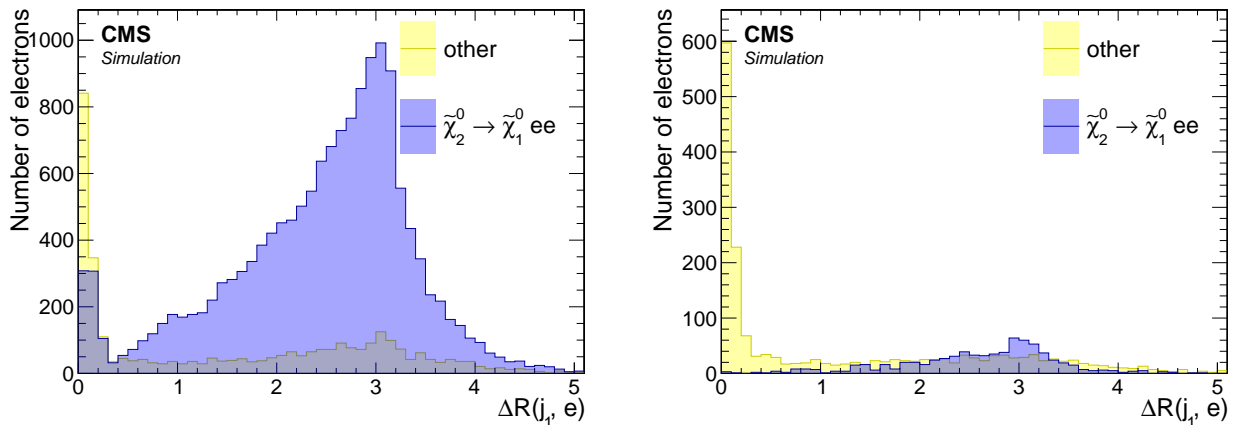


Figure 7.15: Spatial separation between reconstructed electrons with loose ID and the leading jet $\Delta R(j_1, e)$ for $\Delta m = 5.63 \text{ GeV}$ (left) and $\Delta m = 1.92 \text{ GeV}$ (right).

There are two obvious features we can point to in these plots. We have explored the first

feature already in 7.5, namely, that probing lower Δm requires access to low p_T leptons, and since we are limited by a lower threshold of $p_T \geq 5$ GeV on the electrons, that results in lower signal acceptance as can be seen by the difference between the high Δm and the low one. The second interesting feature that we can see, is that our signal electrons are located mainly outside of the leading jet. That is because the leading jet is usually an ISR jet which boosts the $\tilde{\chi}_2^0 \tilde{\chi}_1^0$ system away from it (back-to-back). We therefore make a cut $\Delta R(j_1, e) > 0.4$.

Next we turn to the p_T distributions. First we apply the previous cut of $\Delta R(j_1, e) > 0.4$. As we've already seen in 7.5.4.1, the p_T distribution depends strongly on Δm . Even though the distributions in 7.5.4.1 were plotted using generator level muons, the electrons distributions follow the same trend. We therefore need to make a choice about which Δm to favor, i.e., which Δm we want to be more sensitive to, and we choose the lower Δm case. Nonetheless we compare the two choices in 7.16.

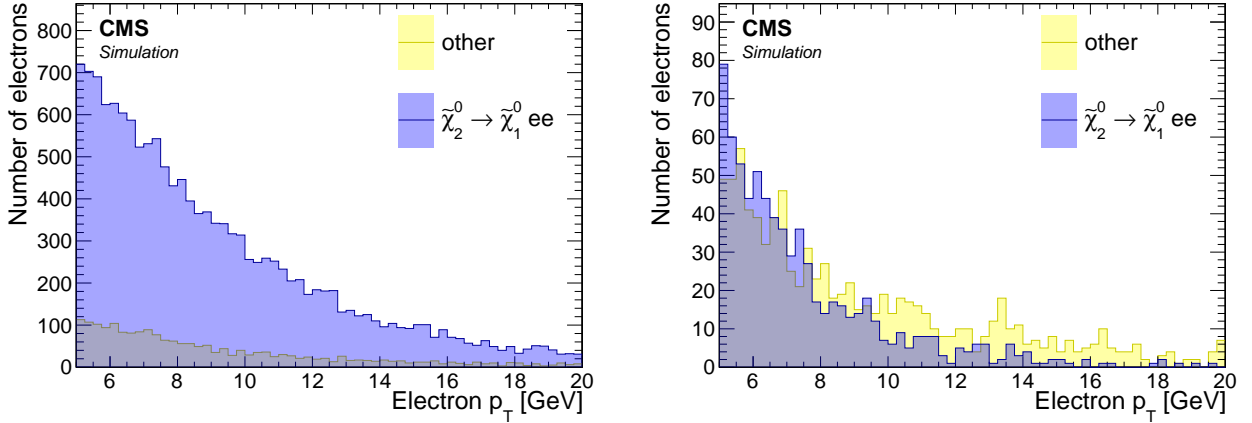


Figure 7.16: p_T distribution of reconstructed electrons with loose ID for $\Delta m = 5.63$ GeV (left) and $\Delta m = 1.92$ GeV (right). Cut of $\Delta R(j_1, e) > 0.4$ applied.

We can see, as expected, that the p_T distribution of the electrons fall more rapidly for the low Δm case. We observe that there are hardly any electrons surviving above 15 GeV, and therefore we choose to make a cut of $p_T < 15$ GeV.

It interesting to look at the η distribution, as seen in 7.17, after the previous cuts to get a better sense of where most of the non-signal electrons are still coming from.

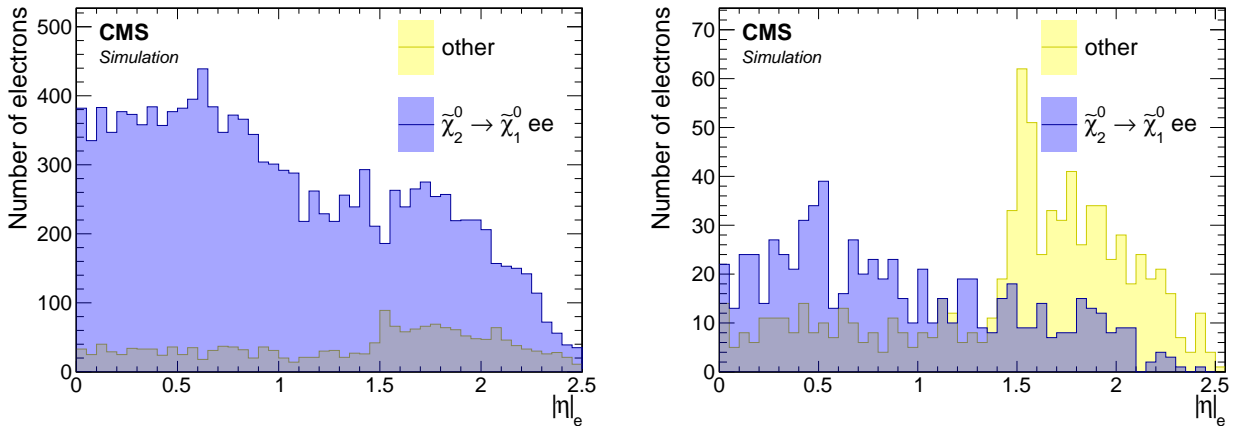


Figure 7.17: $|\eta|$ distribution of reconstructed electrons with loose ID for $\Delta m = 5.63$ GeV (left) and $\Delta m = 1.92$ GeV (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15$ GeV are applied.

In the case of $\Delta m = 1.92$ GeV, we can clearly see how worse the endcaps of the Electromagnetic Calorimeter (ECAL) are performing in comparison with the barrel ($|\eta| < 1.48$). The

transition is clearly visible through a sharp drop in purity at the transition. It is worse for low- p_T electrons than higher- p_T ones.

We would like to see if requiring a tighter working point for the electron-identification is beneficial. The working point used in the previous distributions is loose. We turn now to check the effects of requiring either a medium working point, or a tight one. We plot two bins labeled *fail* and *pass*, which correspond to whether the electron passes or fails the identification criteria of a medium or tight working points. These can be seen in 7.18.

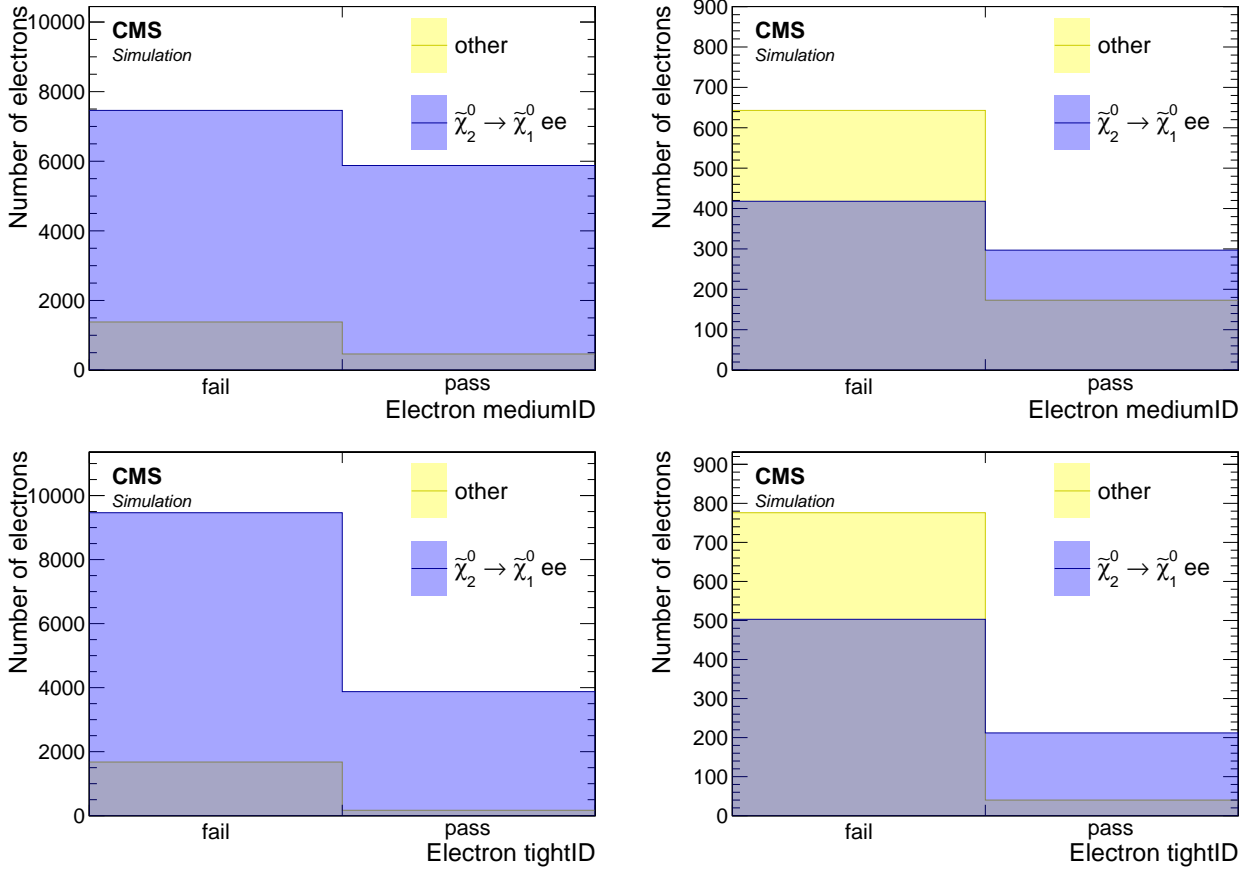


Figure 7.18: Medium (top) and tight (bottom) ID working points distributions of reconstructed electrons for $\Delta m = 5.63$ GeV (left) and $\Delta m = 1.92$ GeV (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15$ GeV are applied.

Selecting a medium or a tight working point is equivalent to choosing the relevant right *pass* bin (top for medium, bottom for tight), and rejecting the electrons on the left *fail* bin. We see that although we reject considerable amount of non-signal electrons in the low Δm case by picking either a medium or tight working points, we also loose quite a lot of signal electrons as well. In other words, these selections are very not efficient and will result in low signal acceptance. We therefore decide to use a loose working point for the electrons. We will see that we can still purify the electron selection by relying on isolation instead. We fully discuss and describe our jet-isolation in 7.7.7, but here for the sake of completeness we look at its effect on the purity of the electrons. We compare our custom jet-isolation to the standard definition of lepton isolation, which does not take into account the possibility that two electrons can be produced close to each other (small ΔR), as is the case in our signal. Comparison of the isolation distributions are seen in 7.19.

We observe that the standard lepton isolation does not perform well in terms of efficiency for both Δm cases. In contrast, the custom jet-isolation is performing very well in terms of signal electron efficiency while successfully rejecting considerable amount of non-signal elec-

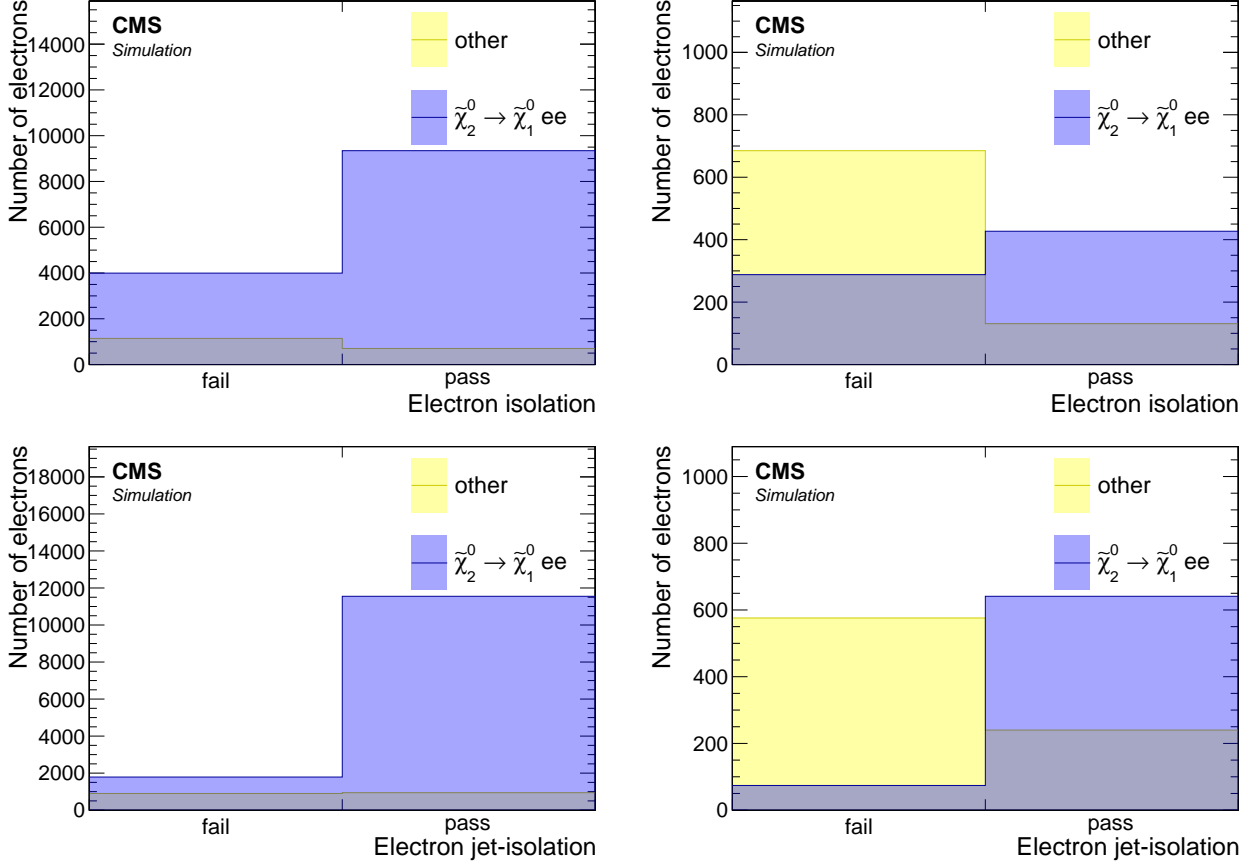


Figure 7.19: Standard isolation (top) and custom jet-isolation (bottom) distributions of reconstructed electrons with loose ID for $\Delta m = 5.63$ GeV (left) and $\Delta m = 1.92$ GeV (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15$ GeV are applied.

trons, resulting in a purer sample of electrons. We therefore conclude that the choice of the custom jet-isolation is favorable. We also look at how the η distribution is affected by this choice at 7.20, and that will conclude our selection of the electrons.

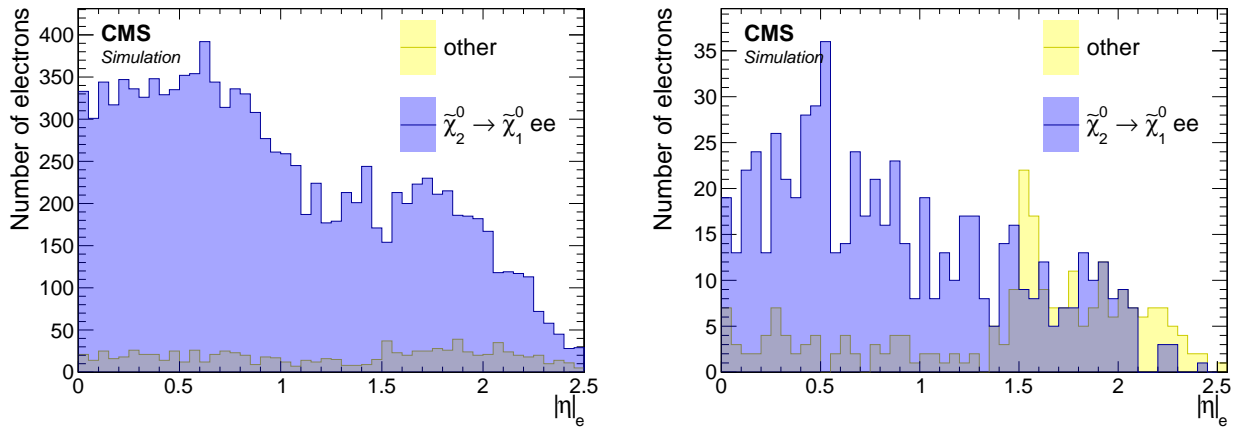


Figure 7.20: $|\eta|$ distribution of reconstructed electrons with loose ID passing jet-isolation for $\Delta m = 5.63$ GeV (left) and $\Delta m = 1.92$ GeV (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15$ GeV are applied.

When we compare distributions 7.20 with 7.17, we can see that our custom jet-isolation has done a good job purifying the electrons selection even further while being efficient in retaining the signal electrons.

We summarize this section with the full selection of the analysis electrons:

- $5 < p_T < 15 \text{ GeV}$
- $|\eta| < 2.5$
- $\Delta R(j_1, e) > 0.4$
- loose ID working point
- pass jet-isolation

7.7.2 Muons

In contrast to the electrons, we do not have an initial reconstruction p_T threshold of 5 GeV. Therefore, we want to explore the possibility of lowering the p_T threshold as much as possible. This has been motivated in section 7.5.4.1, where we saw that the lower Δm we want to probe, the lower p_T threshold we have to allow. Like in the electron case, the initial working point choice for reconstructed muon is loose (see 6). We follow a similar procedure to the electrons case. The first distribution we look at in regards to the muons is their spatial separation from the leading jet in the event, $\Delta R(j_1, \mu)$. We have seen in 7.7 that the muon endcaps are capable of reconstructing muons with $p_T < 3$ GeV while the barrel cannot. It therefore makes sense to look at a split view of barrel and endcaps for the following distributions at 7.21.

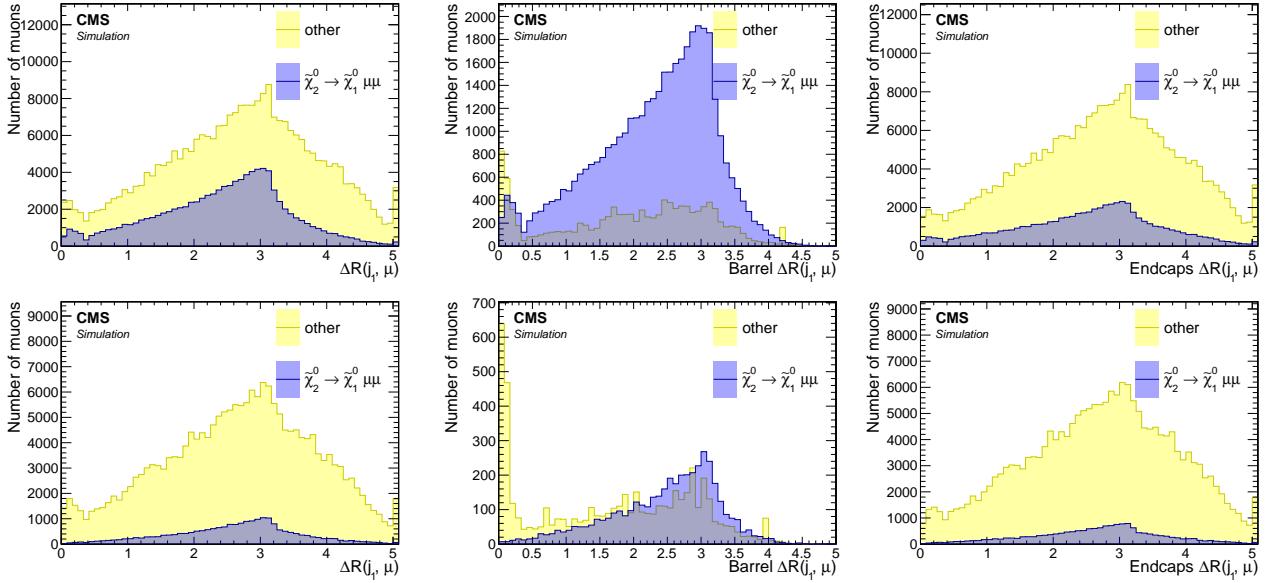


Figure 7.21: Spatial separation between reconstructed muons with loose ID and the leading jet $\Delta R(j_1, \mu)$ for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive case (left), barrel (middle) and endcaps (right).

Since the muons in the endcaps have lower p_T than the muons in the barrel, which is only able to reconstruct muons with $p_T > 3$ GeV, the purity in the endcaps is much lower than the purity in the barrel, and the selection we are constructing here attempts to purify the muons further. Just as in the electrons case, we select muons with $\Delta R(j_1, \mu) > 0.4$, and that selection will apply for the rest of the section.

Next we turn into the p_T distributions. We apply the previous cut of $\Delta R(j_1, \mu) > 0.4$. As we've already seen in 7.5.4.1, the p_T distribution depends strongly on Δm , and we try to favor the low Δm acceptance in order to be more sensitive to it. The p_T distributions we see in 7.22 suggest a cut identical to the electron case of $p_T < 15$ GeV. It is worth mentioning that the p_T of the muons are fed into the training of the Boosted Decision Tree (BDT) for further refinement, and therefore the exact value is being determined here quite loosely. The actual maximum value of the p_T of the muons will depend on the BDT cut being used to define the signal region.

We can see the feature discussed earlier, whereby the endcaps being able to reconstruct muons with lower p_T , and therefore has worse purity than the barrel, being reiterated here. It must be stressed that worse purity is due to a much higher efficiency, and therefore, as long as we can purify it further, is not necessarily a bad thing. We see however, that the bulk of the non-signal muons populate the region of $p_T < 2$ GeV, and the ratio of signal muons to non-signal muons is very low in that region. We therefore make an additional cut of $p_T > 2$ GeV.

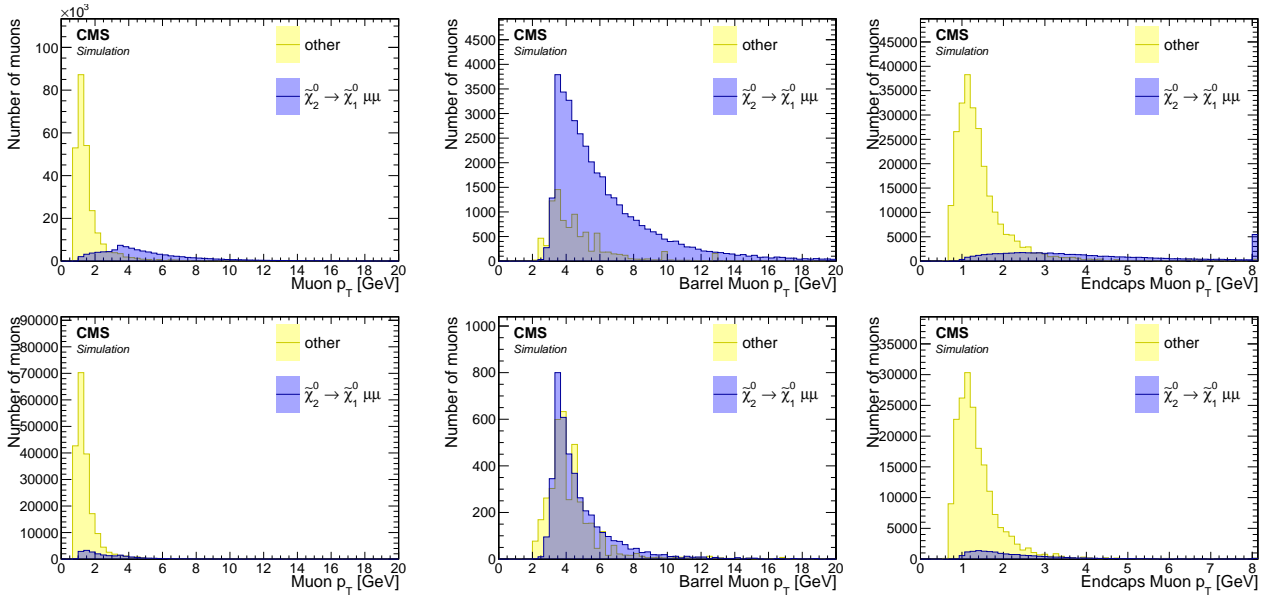


Figure 7.22: Reconstructed muons p_T distribution with loose ID for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$ and $p_T < 15$ GeV are applied.

Another way of looking at the effect of this cut is by looking at the $|\eta|$ distribution before and after the p_T cut, which can be seen in 7.23.

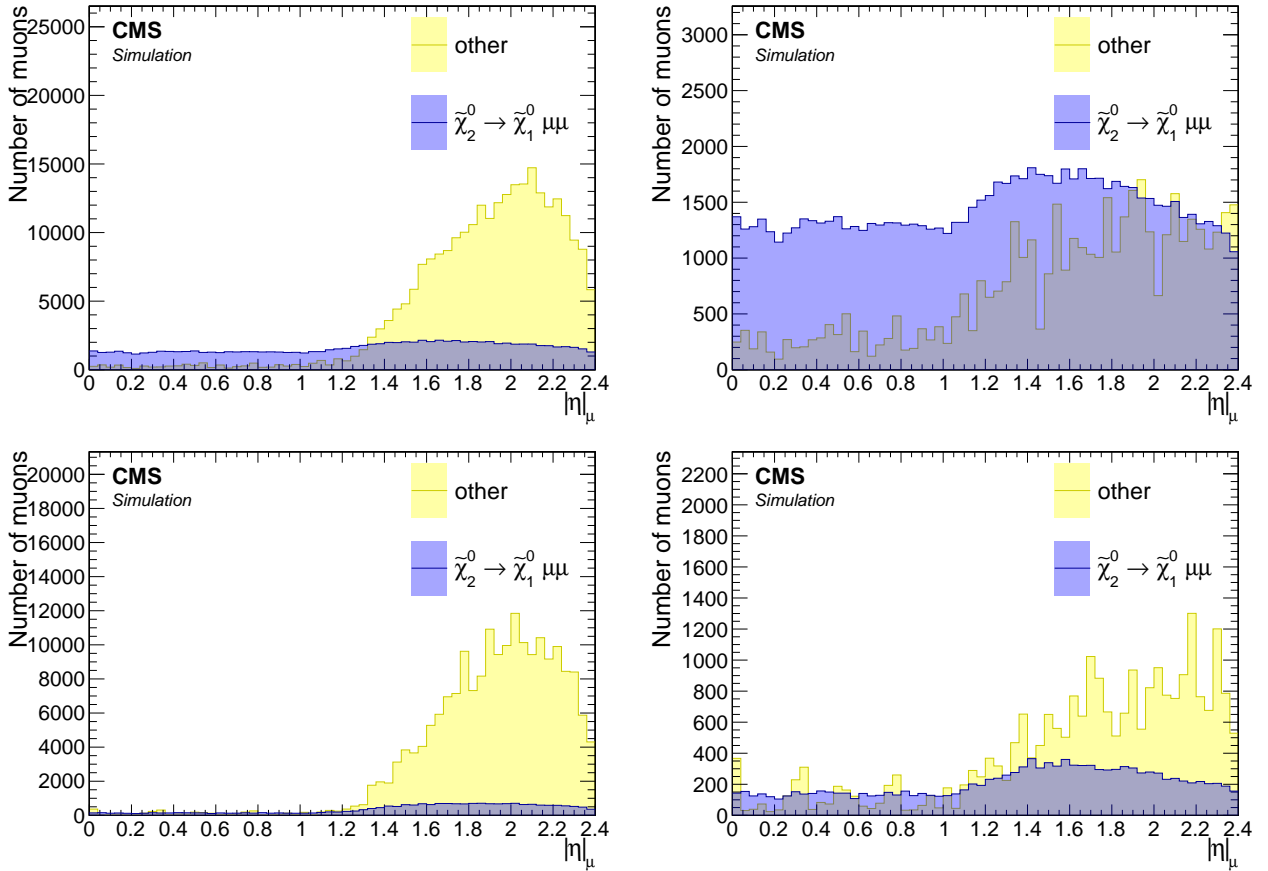


Figure 7.23: $|\eta|$ distribution of reconstructed muons with loose ID for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) without (left) and with (right) $p_T > 2$ GeV cut. Cut of $\Delta R(j_1, \mu) > 0.4$ is also applied.

We would like to see if requiring a tighter working point for the muon-identification is beneficial. The working point used in the previous distributions is loose. We turn now to check the effects of requiring either a medium working point, or a tight one. We plot two bins labeled *fail* and *pass*, which correspond to whether the muon passes or fails the identification criteria of a medium or tight working points.

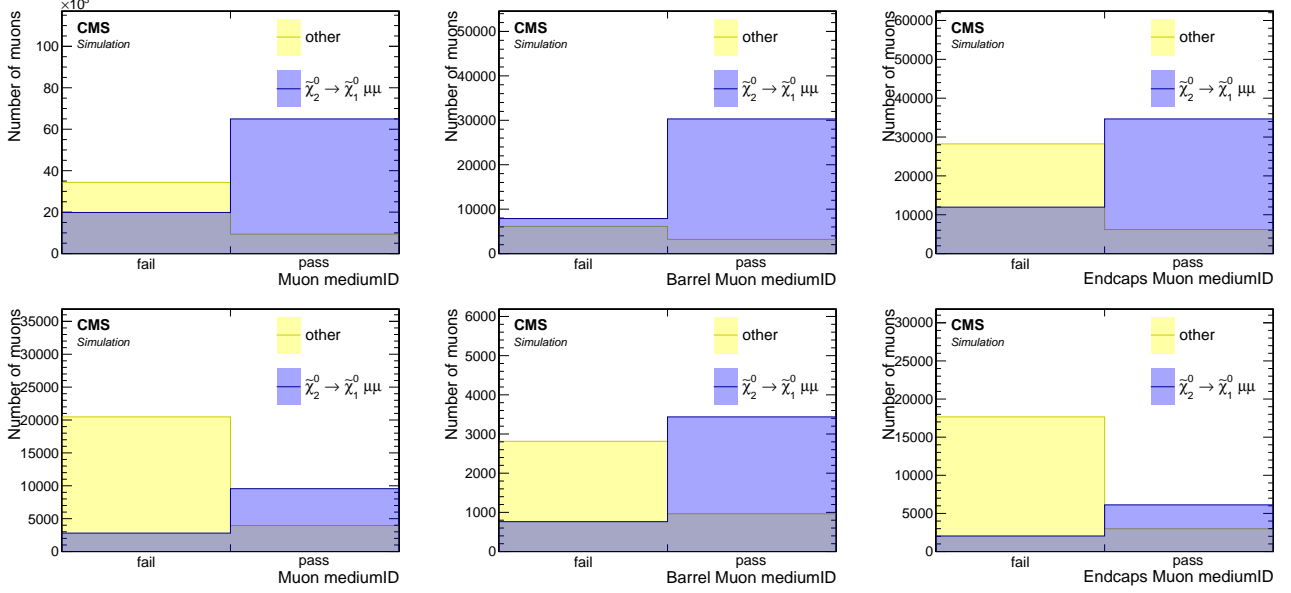


Figure 7.24: Medium ID working point distributions of reconstructed muons for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive p_T case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$, $p_T > 2$ GeV and $p_T < 15$ GeV are applied.

When we compare the medium working point in 7.24 to the tight working point in 7.25 we can see that the medium working point purifies the muons quite a lot and is very beneficial. However, when we look at the tight working point, we observe that we lose quite a lot of our wanted signal-muons without a significant gain in purity. We therefore choose to use the medium ID working point.

Our custom jet-isolation, which is described fully in 7.7.7, was devised mainly to reject SM background while retaining signal. In the electron case, we have seen in 7.19 that it did a great job in also purifying the electron selection and replaced the need of requiring a tighter identification working point. In the case of the muons, we do rely on the a medium working point to perform this task, but we would like to also see the effects of the isolation on our signal muons. We see in 7.26 that we pay a small price by requiring the isolation, but as will be seen in 7.7.7, we increase the sensitivity by rejecting a lot of SM background in the process.

We summaries this section with the full selection of the analysis muons:

- $2 < p_T < 15$ GeV
- $|\eta| < 2.4$
- $\Delta R(j_1, \mu) > 0.4$
- medium ID working point
- pass jet-isolation

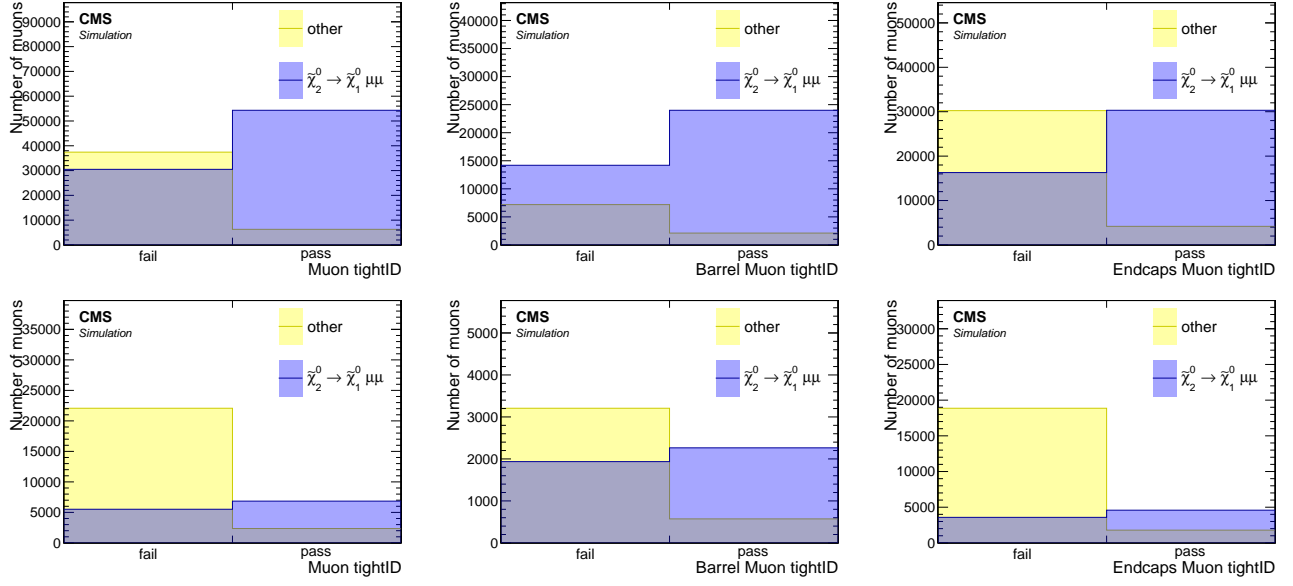


Figure 7.25: Tight ID working point distributions of reconstructed muons for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive p_T case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$, $p_T > 2$ GeV and $p_T < 15$ GeV are applied.

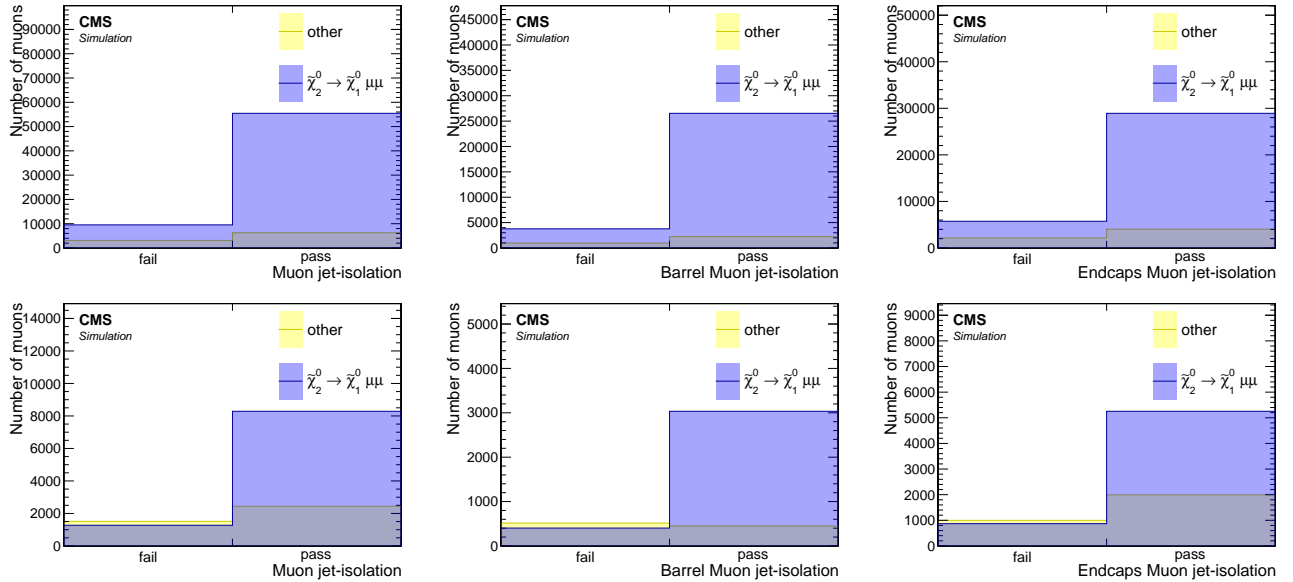


Figure 7.26: Jet-isolation distributions of reconstructed muons with medium ID for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive p_T case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$, $p_T > 2$ GeV and $p_T < 15$ GeV are applied.

7.7.3 Scale factors

In 7.7.1 and 7.7.2 we have studied the selection we apply to electrons and muons respectively. We have therewith made a choice about identification working point. We have used simulation exclusively to draw conclusions about the identification efficiency of the leptons. If we rely on Monte Carlo (MC) simulation, that will produce large systematic errors due to imperfections in modeling both the data and the detector response. We therefore wish to measure the identification efficiency in data, in order to correct the simulation's potentially false efficiency rate. *Efficiency* is defined as how probable it is to reconstruct or identify a lepton. For a lepton ℓ the identification efficiency is defined as:

$$\varepsilon_{\ell}^{\text{ID}} = \frac{N_{\ell}(\text{ID})}{N_{\ell}(\text{produced})} \quad (7.1)$$

In MC simulation, the number of leptons produced is the same as the number of leptons generated. In data, one must measure it using a data-driven method. Once the efficiencies have been measured both in simulation and in data, a correction factor named Scale Factor (SF) can be applied to the simulation in order to correct for any discrepancies that might arise. The scale factors are defined as the ratio between the efficiency in data to the efficiency in simulation:

$$\text{SF}_{\ell}^{\text{ID}} = \frac{\varepsilon_{\ell}^{\text{ID,Data}}}{\varepsilon_{\ell}^{\text{ID,MC}}}, \quad (7.2)$$

dropping the superscript ID we get:

$$\text{SF}_{\ell} = \frac{\varepsilon_{\ell}^{\text{Data}}}{\varepsilon_{\ell}^{\text{MC}}}. \quad (7.3)$$

Once the relevant SF have been determined, they are applied for every lepton passing the object selection in the event. The scale factors for loose-ID electrons in our p_{T} range have been measured centrally by the relevant working group and are applied to the selected electrons. As was determined in 7.7.2, our analysis signal muon's lower p_{T} threshold is 2 GeV which is low. Scale factors for medium ID leptons with $p_{\text{T}} \geq 2$ GeV were computed centrally by the Muon Physics Object Group (POG). The scale factors, however, while matching our muons' p_{T} range and identification working point, were computed by requiring $\Delta R > 0.5$ between the muons [2, 3]. As we have seen in 7.5.4.3, one of the drivers of the sensitivity is the region of $\Delta R < 0.5$. We would therefore like to validate the scale factors in that region. We would like to show that the efficiencies have no ΔR dependence, and in order to do so, we calculate the efficiencies in different ΔR regions.

In order to measure such efficiencies in data, one must identify desired leptons with low and easily reducible fakes. A widely used method to perform such a data-driven task is the Tag & Probe method. In the Tag & Probe method we examine a mass resonance such as Z, J/ ψ or Y to select particles of the desired type, and probe the efficiency of a particular selection criterion on those particles. The mass resonance will then decay into two same-flavor opposite charged pair of leptons and will form a peak on top of a background. Since we are interested in measuring the efficiency of low p_{T} muons, we choose to look at dimuon events around the J/ ψ mass window. In a dimuon event, we describe one muon as a 'tag' and the other as a 'probe'. The tag muon is selected with a very tight selection which results in very high certainty that the object corresponds to a real muon produced. The probe is given a loose selection, but since it is constrained to be consistent with a product of a J/ ψ , it is almost certain that it originates from a real muon too. Since the shape of the J/ ψ is a peak over a background, the background

is easily removed by a fit. The probe is then subjected to cuts or constraints which are used to measure a particular efficiency. As stated, in this study we want to show that the efficiency to identify a muon with medium ID working point from a track has no ΔR dependance. Therefore, the efficiency we are looking for is defined as:

$$\varepsilon_{\mu}^{\text{ID}} = \frac{N_{\mu}^{\text{ID}}}{N_t}. \quad (7.4)$$

The probe we are using in the denominator is a track passing a loose selection, while the probe track in the numerator is required to match a medium ID working point muon. The number of objects passing a selection is determined by a fit to data and MC, to measure the corresponding efficiencies in data and MC.

This study is done for year 2016. For the MC, we are using the 2016 samples listed in 7.6.1. For data, we are using a single electron trigger, in order for the tagged muon to be independent from the triggered object. The data set is measured to correspond to 36.02 fb^{-1} using the BRIL Work Suite [4]. The following trigger paths are used:

- HLT_Ele27_WPTight_Gsf_v*,
- HLT_Ele27_eta2p1_WPLoose_Gsf_v*,
- HLT_Ele32_WPTight_Gsf_v*,
- HLT_Ele35_WPTight_Gsf_v*.

We then select an offline loose ID electron with $p_T > 27 \text{ GeV}$. The requirements to select a tag & probe pair are defined in table 7.7.3.

Table 7.3: Selection criteria for Tags and Probes

Tag	Probe
medium ID muon	isolated track
$p_T \geq 5 \text{ GeV}$	$2 \leq p_T \leq 20 \text{ GeV}$ ($p_T \geq 3 \text{ GeV}$ for barrel)
$ \eta < 2.4$	opposite-sign in invariant mass window $[2.5, 3.5] \text{ GeV}$

A fit is then performed in an invariant mass window around the J/ψ window of $[2.5, 3.5] \text{ GeV}$. The signal fit is using a crystal ball function and the continuum is fit with a 6th order polynomial. The fit is repeated twice, where the denominator is done with probe tracks, and the numerator is using medium ID muons that have been matched to said tracks. The ΔR range has been split into 3, and $|\eta|$ of the muons has been split into barrel ($|\eta| < 1.2$) and endcaps ($1.2 < |\eta| < 2.4$). Simulation fits are shown in 7.27 for barrel, and 7.28 for endcaps. Data fits are shown in 7.29 for barrel, and 7.30 for endcaps.

The efficiencies and corresponding scale factors can be seen in 7.31. The scale factors are statistically consistent with unity, and show no discernible ΔR dependence. A similar study has been carried out with simulation and data for 2017 and 2018 in [5] and did not observe a ΔR dependence either. The recommendation from the POG as a results of these studies are to use the calculated scale factors provided by them with an additional systematic of 1% for muons with $p_T < 20 \text{ GeV}$.

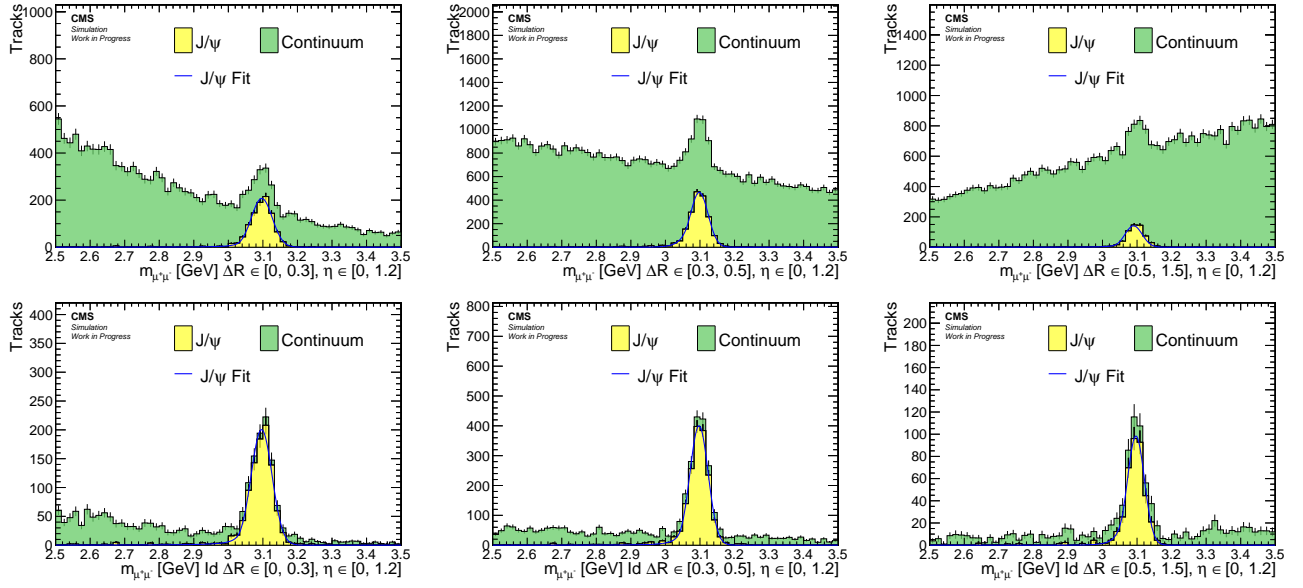


Figure 7.27: Simulation barrel muons fits for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right)

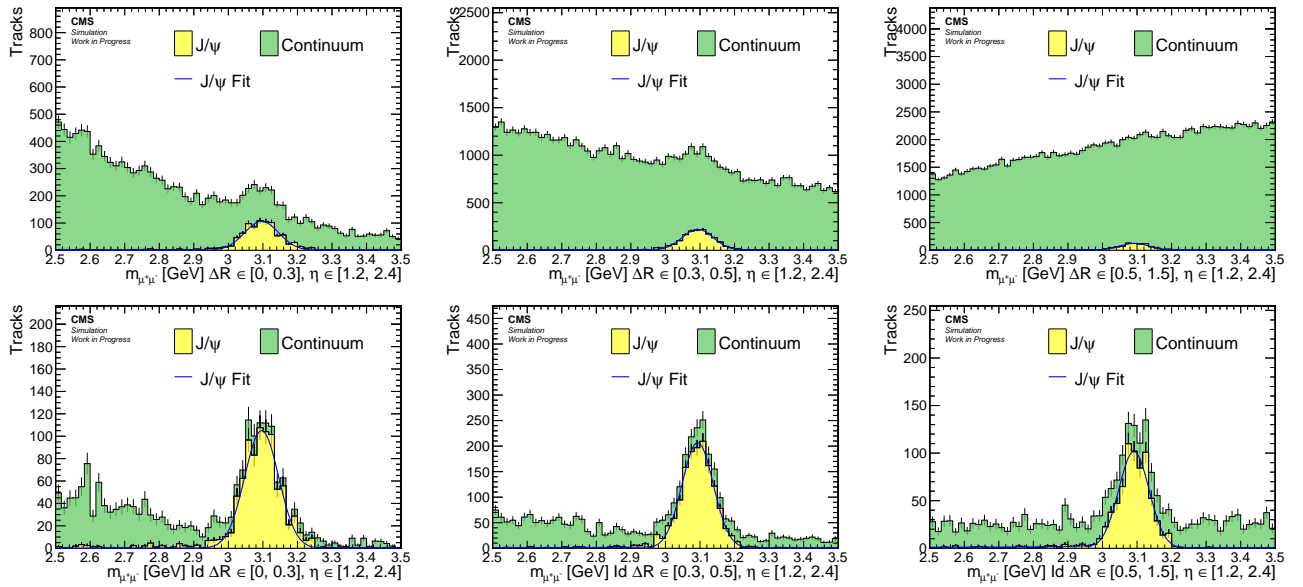


Figure 7.28: Simulation endcaps muons fits for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right)

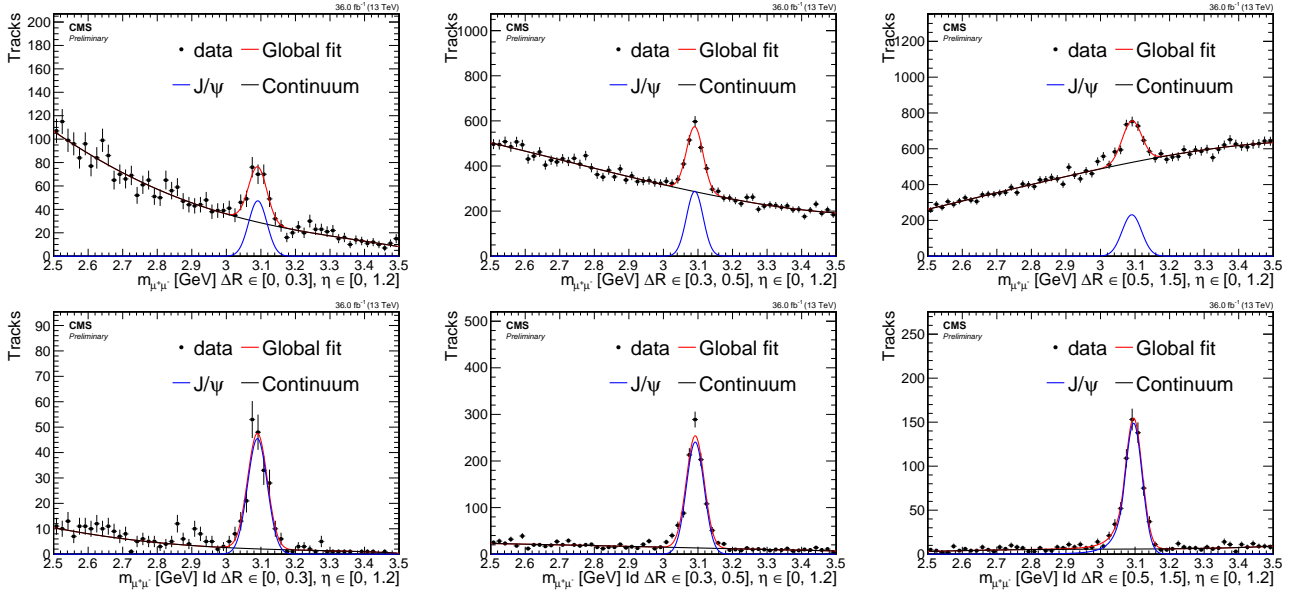


Figure 7.29: Data barrel muons fits for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right)

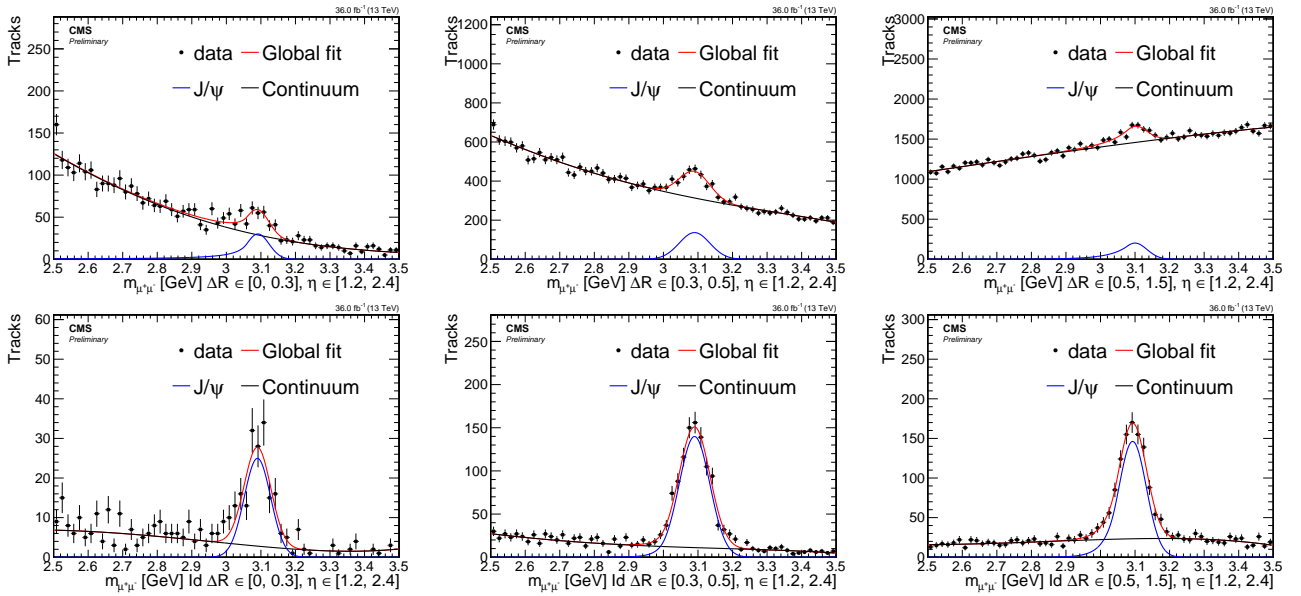


Figure 7.30: Data endcaps muons fits for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right)

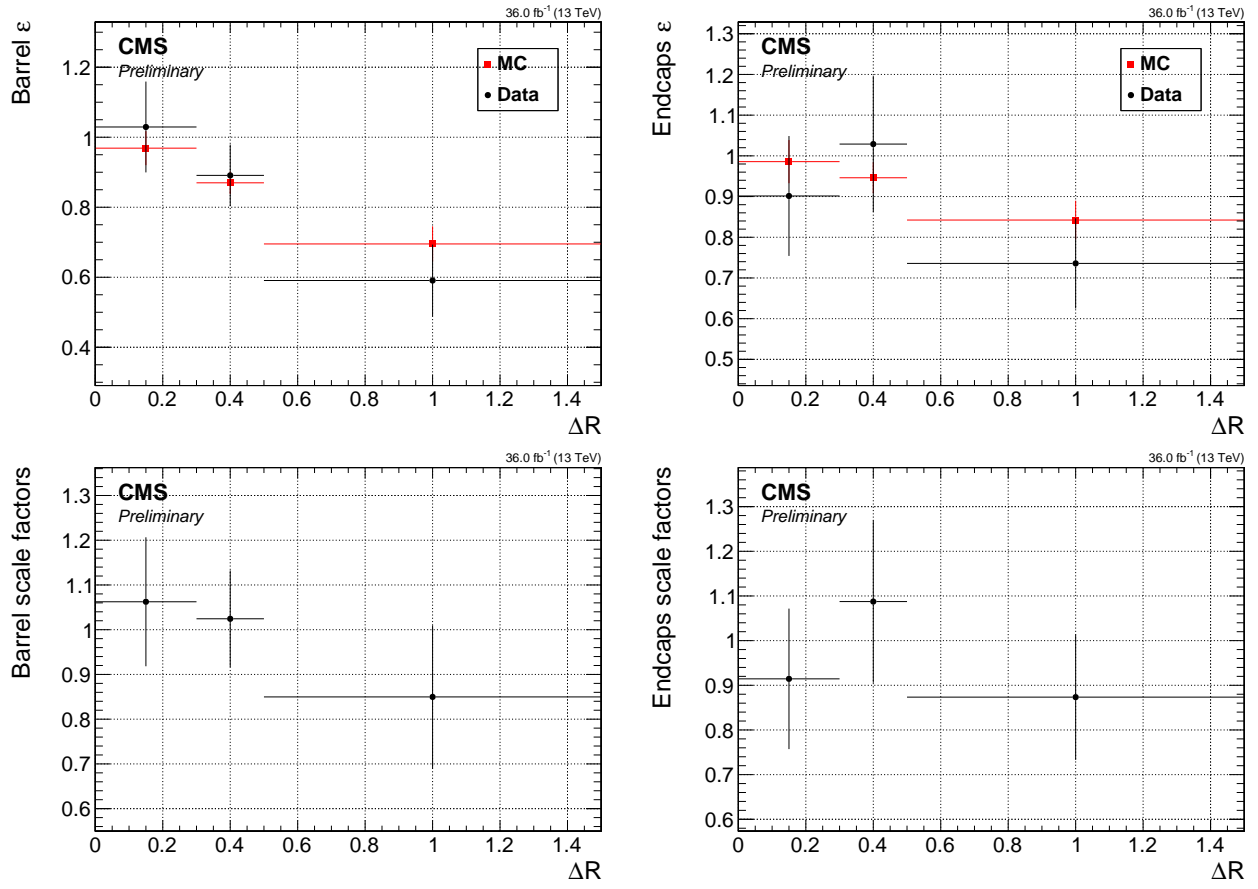


Figure 7.31: Efficiencies (top) and scale factors (bottom) for barrel muons (left) and endcaps muons (right).

7.7.4 Missing transverse energy

The importance of measuring the missing transverse momentum (or energy) in this analysis has been discussed in 7.5.1. We concluded that the missing transverse momentum is an essential ingredient in our analysis strategy. It is used in order to trigger events online, as well as in our offline event level selection in order to boost sensitivity. Two standard measures of the momentum imbalance in the events are \vec{E}_T^{miss} (or equivalently by a different symbol \vec{p}_T^{miss}) and \vec{H}_T^{miss} . Although \vec{E}_T^{miss} and \vec{p}_T^{miss} have different symbols, and referred to by different names (missing transverse energy and missing transverse momentum respectively), they are defined in the same way and are used interchangeably. Mathematically speaking, \vec{E}_T^{miss} is defined in the following way:

$$\vec{E}_T^{\text{miss}} = \vec{p}_T^{\text{miss}} = - \sum_i \vec{p}_T(i) \quad (7.5)$$

where the summation is done on all particle flow candidates. Therefore, it is a measure of particles that escape detection, such as weakly interacting particles. The missing transverse energy is highly sensitive to any mismeasurements of the visible particles, as well as additional energy deposits from Pile-Up (PU), detector noise etc. Therefore, this observable is further corrected to mitigate PU effects, as well as jet energy response. For those corrections, jets with p_T greater than 10 GeV are being considered. Full details of the corrections are given in [6].

An alternative measurement to the missing transverse momentum is \vec{H}_T^{miss} , sometimes referred to as *missing hardronic activity*. Rather than looking at all particle flow candidates in the sum, it takes into account only jets with p_T greater than 30 GeV with $|\eta| < 5$:

$$\vec{H}_T^{\text{miss}} = - \sum_i^{\text{jets}} \vec{p}_T(i) \quad (7.6)$$

In this analysis we favor using the observable \vec{H}_T^{miss} over \vec{E}_T^{miss} , since, as we see in our definition of the jet isolation in 7.7.7, we use jets with p_T greater than 30 GeV, and we keep the range of $p_T \in [15, 30]$ GeV of jets to define a side band which we then use for our data-driven background estimation method in 7.10.2.1. For both observables, \vec{E}_T^{miss} and \vec{H}_T^{miss} , we can define the equivalent scalar quantities, E_T^{miss} and H_T^{miss} respectively, by taking the magnitude of their vectorial counterpart.

7.7.5 Jets

Jets reconstruction and identification is described in 6.4. Jets used in the analysis are reconstructed based on a clustering of the Particle Flow (PF) candidates using FASTJET with the anti- k_T algorithm [7] with the size parameter 0.4. Tagging of b quark jets is done using the so-called Combined Secondary Vertex (CSV) algorithm based on a multivariate technique DEEPCSV with a medium working point. Each jet is required to have $p_T > 30$ GeV and $|\eta| < 2.4$.

7.7.6 Tracks and multivariate selection

The leptons $\ell^+\ell^-$ produced in our decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$ tend to have mostly very low transverse momentum p_T . We have seen in 6 that the identification and reconstruction of the muons get worst with lower p_T . Therefore, the aim of the exclusive track category is to try and regain lost leptons that didn't make the reconstruction or identification process. Since the tracking efficiency at our p_T ranges is well above 99%, as we've seen in 6.1, we can try and pick up the track that corresponds to our missing lepton by applying our knowledge about the event kinematics specific to our signal. We are therefore, not enhancing the lepton identification in the general case, but identifying missed tracks in our specific signal signature.

Since each event contains many tracks, we need a method to identify in the signal event, which track corresponds to our miss-identified lepton. In order to achieve this goal we train a Boosted Decision Tree (BDT). We train 4 BDTs which correspond to the lepton flavor, Muons or Electrons, and for each, we train separately for phase 0 (2016) and phase 1 (2017-2018) of the tracker. All BDTs use the same structure of 200 trees with a maximum depth of 3, with the TMVA package [8]. The BDT training is performed with AdaBoost and GiniIndex separation. We are taking all other values as the defaults set by the TMVA package.

For the training we take tracks from a pool of our privately produced FASTSIM signal simulations which were listed in 7.6.2. Of those, we are selecting the full range of simulated higgsino parameter μ (or the mass of $\tilde{\chi}_1^\pm$ in case of phase 1), but only the range of Δm we want to be most sensitive to. In phase 0, we select $\Delta m^0 \in [0.3, 4.3]$ GeV and $\mu \in [100 - 130]$ GeV. In phase 1 we select $\Delta m^\pm \in [0.3 - 4.6]$ GeV and $\mu \in [100 - 500]$ GeV. Those signal events we then split into signal tracks, i.e., tracks originating from the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$ and been matched to the missing generated lepton in the generator level particles collection, and background tracks which do not match our wanted leptons. Therefore, our BDTs are useful to reject in-signal background of unwanted tracks. The samples for muons contain 9408 (10964) signal tracks and 99996 (151380) background tracks for phase 0 (phase 1). For electrons the samples contain 2364 (2288) signal tracks and 104065 (159713) background tracks for phase 0 (phase 1). The training samples are then tested against the test samples of equal size. The distributions of the testing samples overlay on the training samples are seen in 7.33.

A pre-selection is applied to all tracks in the collection obtained by the standard track reconstruction sequences. The pre-selection ensures a set of properly-reconstructed, isolated and prompt tracks, whose trajectories pass through the region nearby the primary vertex with the largest sum of charged-tracks, jets and missing energy values (PV):

- $p_T > 1.9$ GeV
- $|\eta| < 2.4$
- track $\text{iso}_{\text{rel}} < 0.1$, using $\Delta R(\text{track}, \text{other tracks}) < 0.3$
- $d_{xy}(\text{track}, \text{PV}) < 0.02$ cm w.r.t the PV
- $d_z(\text{track}, \text{PV}) < 0.02$ cm w.r.t the PV
- no match to an electron or muon within a cone of size 0.01

For the training we use 10 variables listed in decreasing order of their ranking (in the muon case of phase 0) listen in 7.7.6.

Distribution for the input variables showing the signal tracks in blue, and background tracks in red, are seen in 7.32.

The output of the training of the 4 BDTs can be then seen in 7.33. The testing distributions are laid on top of the training sample of the BDT. No obvious over-training is seen. The ROC

Table 7.4: Track BDT input variables

Rank	Variable	Description
1	$\Delta R(t, \ell)$	t is the track and ℓ the lepton
2	$ \Delta\eta(t, \ell) $	
3	$p_T(\ell)$	
4	$ \Delta\phi(t, \vec{H}_T^{\text{miss}}) $	j_1 is the leading jet
5	$ \Delta\eta(t, j_1) $	
6	$ \Delta\phi(t, \ell) $	
7	$ \eta(t) $	invariant mass
8	$ \eta(\ell) $	
9	$\Delta R(\ell, j_1)$	
10	$m_{t\ell}$	

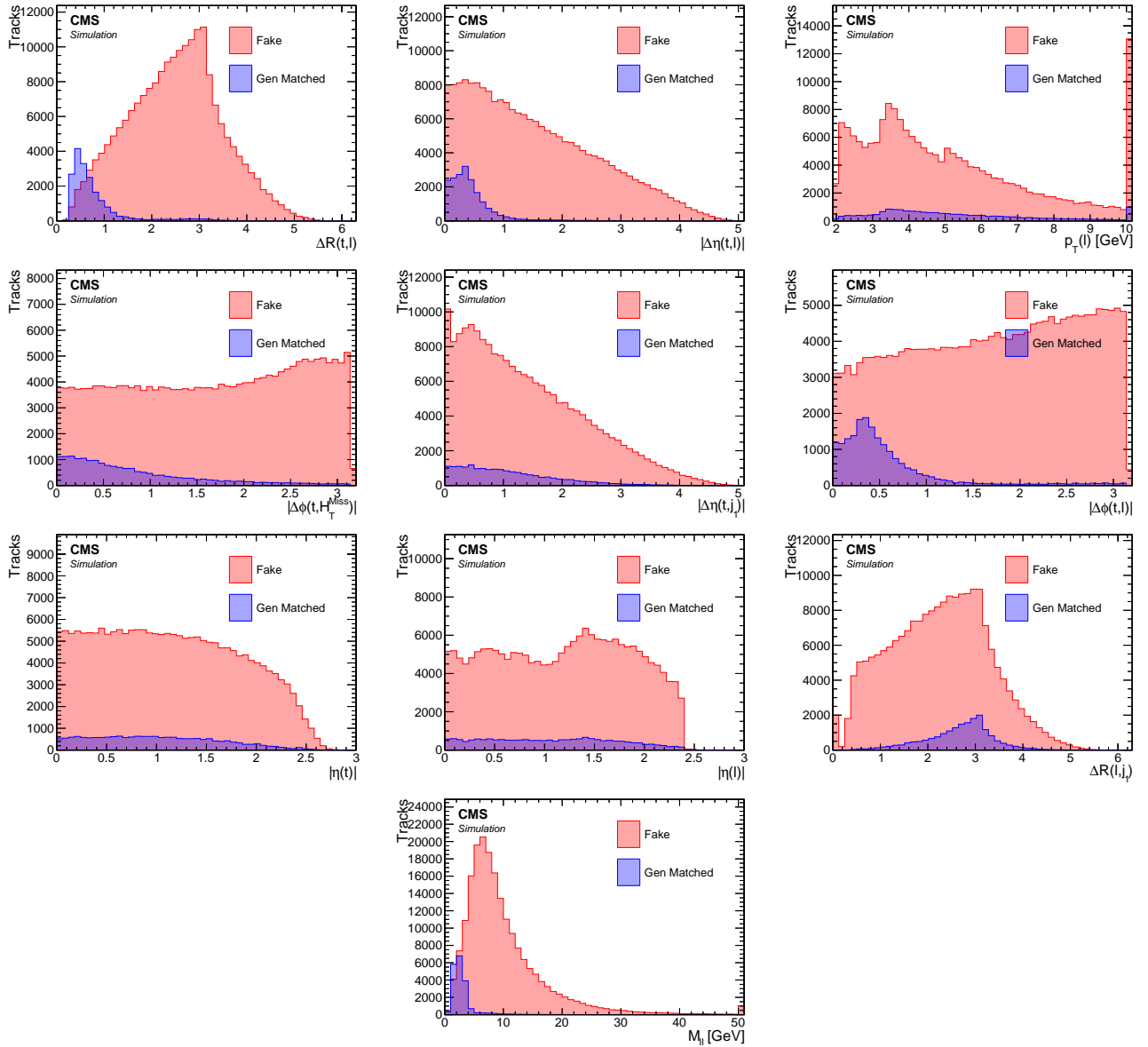


Figure 7.32: Distributions of the inputs for training for the track BDT in the Muon exclusive track category.

curves are then plotted in 7.34. The red point shows the efficiency of the signal and background tracks of the BDT cut which is chosen to be 0.0. We get quite a good separation of signal tracks

from fake tracks, as can be seen by the relatively high signal efficiency of over 90% (86%) for muons (electrons) and background rejection of around 86% (76%) for muons (electrons).

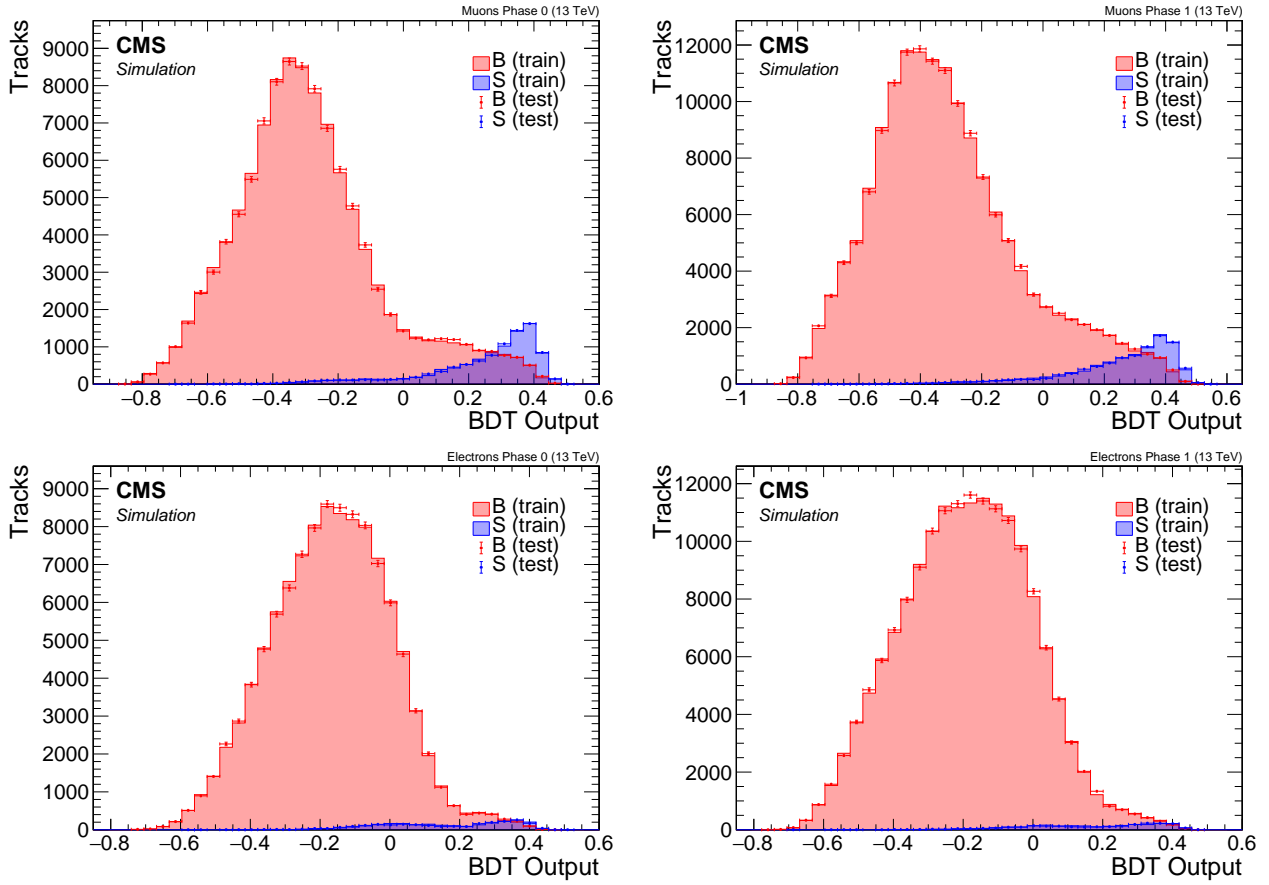


Figure 7.33: Track BDT output plots for Muons (top) and Electrons (bottom) in phase 0 (left) and phase 1 (right). Blue shows signal tracks, while Red are fake tracks. Test sample overlay on top of training sample.

After the training process, we select the track with the maximum BDT score as our missing lepton in the event. We consider only events with a track with a score of greater than 0.0, which corresponds to the red dot in the ROC curves in 7.34.

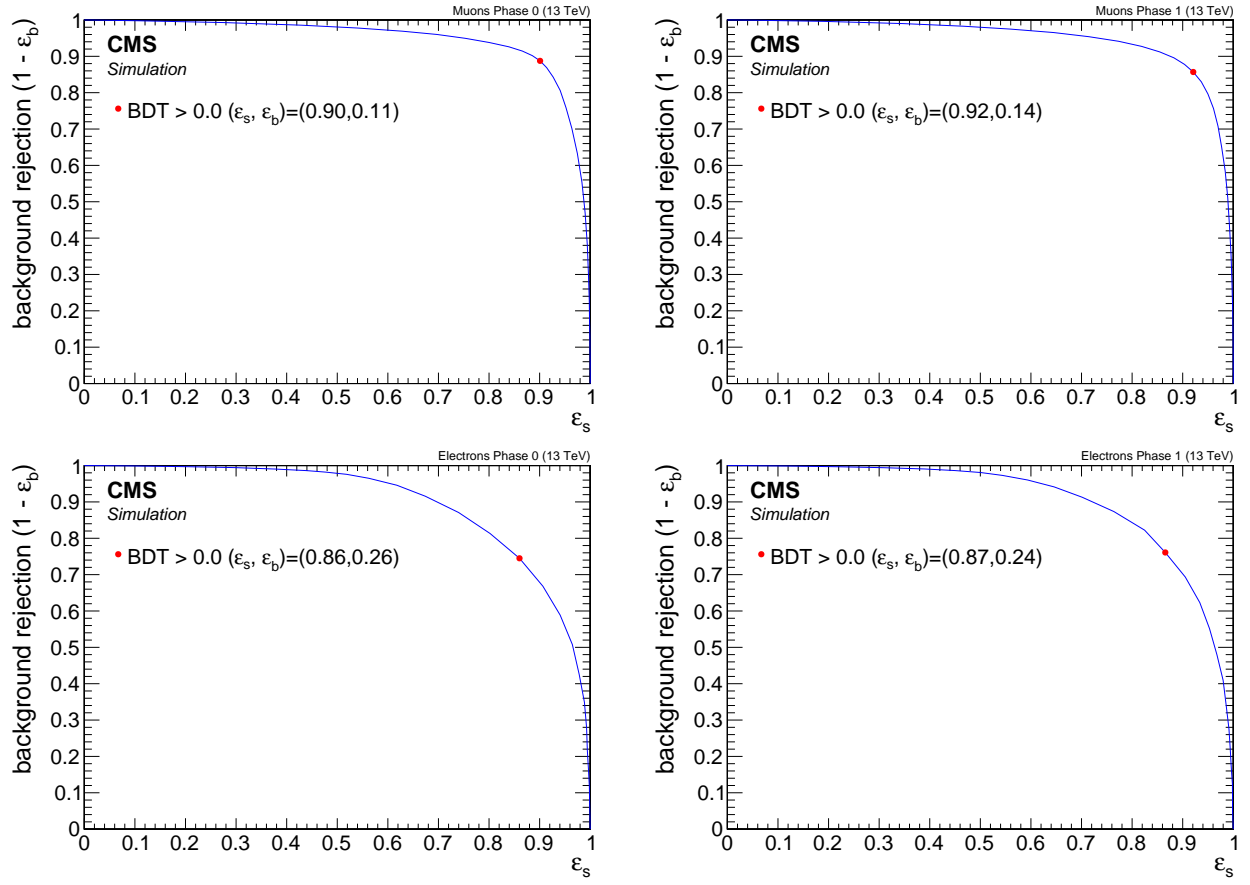


Figure 7.34: Track BDT ROC curves for Muons (top) and Electrons (bottom) in phase 0 (left) and phase 1 (right). Minimum working point showed as a red dot.

7.7.7 Isolation

The leptons that are produced from the neutralino decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$ are typically clean and isolated with very little hadronic activity around them. That is due to the fact that the only jets in the event come from initial state radiation which provides a boost to the produced electoweakinos in the other direction. Therefore, the leptons originating from those electoweakinos will not exist in proximity to said jets. We can exploit this signature in order to discriminate between our signal and background originating from standard model processes in association with jets. At CMS various standard isolation criteria were developed and used. The three most widely used isolation criteria are track isolation [9], relative isolation (RelIso), which was first described in [10], and a modified version refereed to as relative mini-isolation (miniRelIso) described in [11].

Track isolation is defined as the p_T sum of all tracks around a given track (or lepton) in a fixed cone size of 0.3:

$$\text{Track Isolation}_\ell = \frac{\sum_{\text{tracks from PV in } \Delta R < 0.3} p_T}{p_T(\ell)} \quad (7.7)$$

Since we are summing tracks, we take into account only charged particles. Another widely used isolation is the relative isolation which uses a cone size of 0.4 and defined as:

$$\text{RelIso}_\ell = \frac{\sum_{\text{charged hadrons from PV}} p_T + \max \left(0, \sum_{\text{neutral hadrons}} E_T + \sum_{\text{photons}} E_T - 0.5 \cdot \sum_{\text{charged hadrons from PU}} p_T \right)}{p_T(\ell)} \quad (7.8)$$

The last term in the definition is correction due to PU effects. A lepton is said to pass isolation if it is required to have a small RelIso value. The third widely used isolation criterion is the mini-relative isolation miniRelIso, and is a modified version of the relative isolation. The difference is that in the mini-isolation, the cone size depends on the p_T of the lepton in the following way:

$$R = \begin{cases} 0.2 & p_T(\ell) \leq 50 \text{ GeV} \\ \frac{10 \text{ GeV}}{p_T(\ell)} & p_T(\ell) \in (50 \text{ GeV}, 200 \text{ GeV}) \\ 0.05 & p_T(\ell) \geq 200 \text{ GeV} \end{cases} \quad (7.9)$$

The variable size cone allows to recover efficiency when leptons are produced in the decay chain of a boosted object. In such cases, it is likely that, when the boost is large, the lepton overlaps with a jet produced in the same decay chain, failing a standard isolation cut.

The drawback of the standard isolation criteria is that if two leptons are in proximity of each other, as is often the case in our signal, they will spoil each other isolation. We have seen in 7.5.4.3 that in order to gain access to low Δm model-points, we need to also gain access to the $\Delta R < 0.3$ phase-space region. That means that requiring any of the standard isolation criterion will result in rejecting valuable signal events. We therefore propose an alternative isolation criterion that will help retain some of the wanted phase-space while rejecting the majority of the standard model background. Another crucial use for an alternative isolation criterion is background estimation, as is described in 7.10.2.1 Here we describe the steps to construct the alternative *jet isolation*. It is defined for each lepton flavor individually, since the leptons produced are of same flavor. The steps are as follows (introducing parameters p and r):

1. Create a copy set of the jets collection
2. Create a *lepton-corrected* set of the jets by subtracting all leptons within the jet cone of 0.4
3. Clean the set of jets by keeping only jets with positive energy and p_T greater than threshold p
4. Lepton is said to pass *jet-isolation* if it does not lie within a cone of size r from a *lepton-corrected* jet
5. Lepton is said to fail *jet-isolation* for background estimation if it fails *jet-isolation*, and the original jet closest to it has $15 < p_T < 30$ GeV (see 7.10.2.1 for use of such lepton)

The main idea behind the definition of the jet-isolation is that we want to reject leptons with hadronic activity around them, but not losing a lepton that is close to another same-flavor lepton. The process we described introduced two free parameters, namely, the p_T threshold of the lepton-corrected jets responsible for failing a lepton's isolation p , and the cone size r which determines how close a corrected jet is allowed to be to a lepton. In order to choose the thresholds for these parameters, a scan has been performed ranging $p \in [0, 20]$ GeV and $r \in [0.4, 0.6]$. The scan is used to compare different criteria for the optimal values. Since we are also using the isolation for our background estimation, some of the criteria we are interested in stem from that method. The criteria we are interested in are signal efficiency (which we want to be high), background efficiency (which we want to be low), signal contamination in our control-regions (ideally low), jetty-background transfer factor (ideally less than 1), and lastly, the significance, which is computed taking into account transfer factor error on the background (which ideally we want to maximize). To demonstrate the effect these parameter space has on these criteria, we can look at a scan done for the muons using 2016 MC and data.

Table 7.5: Signal Efficiency

	ΔR				
	0.4	0.45	0.5	0.55	0.6
p_T 0	0.38	0.37	0.36	0.35	0.35
1	0.39	0.38	0.37	0.37	0.36
5	0.65	0.64	0.63	0.62	0.60
6	0.71	0.70	0.69	0.67	0.66
7	0.77	0.76	0.74	0.73	0.72
8	0.82	0.82	0.80	0.78	0.77
9	0.87	0.86	0.85	0.84	0.82
10	0.89	0.89	0.87	0.86	0.85
10.5	0.90	0.90	0.89	0.88	0.87
11	0.92	0.92	0.91	0.90	0.89
11.5	0.93	0.92	0.91	0.91	0.90
12	0.94	0.93	0.92	0.91	0.90
12.5	0.94	0.94	0.93	0.92	0.91
13	0.95	0.95	0.94	0.93	0.93
15	0.98	0.98	0.97	0.97	0.97
20	1.00	1.00	1.00	0.99	0.99

We can see from the scan, that the transfer factor of the jetty background estimation method becomes larger with larger p_T and smaller with larger ΔR . We opt for a transfer factor that is less than unity, and therefore we exclude those choices. Taking into account all factors, we make the choice of $(p, r) = (10 \text{ GeV}, 0.6)$ for muons and $(p, r) = (10 \text{ GeV}, 0.5)$ for electrons.

Table 7.6: Background Efficiency

	ΔR				
	0.4	0.45	0.5	0.55	0.6
p_T	0	0.08	0.07	0.06	0.06
	1	0.08	0.07	0.06	0.06
	5	0.12	0.12	0.10	0.09
	6	0.15	0.14	0.12	0.11
	7	0.18	0.16	0.15	0.14
	8	0.20	0.18	0.17	0.17
	9	0.25	0.23	0.19	0.18
	10	0.26	0.25	0.22	0.19
	10.5	0.27	0.24	0.23	0.20
	11	0.29	0.26	0.24	0.22
	11.5	0.28	0.27	0.24	0.23
	12	0.29	0.27	0.26	0.24
	12.5	0.31	0.28	0.26	0.26
	13	0.33	0.29	0.27	0.27
	15	0.36	0.33	0.30	0.29
	20	0.45	0.41	0.39	0.36

Table 7.7: Transfer Factor

	ΔR				
	0.4	0.45	0.5	0.55	0.6
p_T	0	0.19	0.16	0.13	0.13
	1	0.18	0.16	0.14	0.13
	5	0.31	0.30	0.26	0.23
	6	0.43	0.36	0.32	0.30
	7	0.55	0.48	0.44	0.40
	8	0.68	0.58	0.52	0.52
	9	0.83	0.78	0.65	0.58
	10	0.99	0.93	0.76	0.67
	10.5	1.07	0.95	0.85	0.74
	11	1.19	1.10	0.93	0.85
	11.5	1.24	1.19	0.96	0.91
	12	1.34	1.29	1.09	0.99
	12.5	1.55	1.35	1.21	1.10
	13	1.70	1.46	1.27	1.23
	15	2.39	2.17	1.80	1.63
	20	6.12	5.86	4.82	4.13

Table 7.8: Significance $s/\sqrt{b + \epsilon_b^2}$

	ΔR				
	0.4	0.45	0.5	0.55	0.6
0	4.29	6.08	6.13	5.89	5.46
1	4.92	5.18	6.34	5.33	5.84
5	6.44	5.27	6.20	8.63	5.98
6	4.72	5.06	6.22	6.99	7.92
7	4.83	6.55	5.09	5.63	6.28
8	3.80	5.48	4.60	5.24	4.61
9	3.60	4.43	5.66	6.25	4.60
10	3.37	4.08	5.57	4.78	6.23
10.5	3.72	4.03	4.90	4.48	4.17
11	3.05	3.51	4.37	4.98	5.41
11.5	3.21	3.21	3.84	3.54	4.65
12	3.48	3.51	3.80	3.30	3.54
12.5	2.79	3.19	2.82	3.36	4.60
13	3.16	2.68	3.59	6.60	3.50
15	4.46	3.19	3.06	3.64	3.85
20	7.21	1.46	1.60	8.10	2.09

7.8 Trigger

7.9 Event selection

In the following section we detail the event level selection applied in this analysis. We have seen in 7.4 that we refer to three categories in this analysis: dimuon category, and exclusive track category for each lepton flavor. We summarise the preselection in 7.9.1, then the selection that defines each category in 7.9.2, and finally the multivariate selection for each category in 7.9.3.

7.9.1 Baseline selection

We have reviewed some of the base selection that is common to all categories in 7.5.3. We reiterate reasons for these selections here in addition to other event level selections.

- $H_T^{\text{miss}} \geq 220 \text{ GeV}$ and $E_T^{\text{miss}} \geq 140 \text{ GeV}$ cuts intend to boost sensitivity by rejecting SM background, as well as to operate in the acceptance regime of the MET trigger, as described in 7.8. It is especially efficient in rejecting QCD background, since it does not produce real E_T^{miss} . Any E_T^{miss} apparent in QCD is due to jet energy miss measurements. The reason we are cutting harder on H_T^{miss} rather than E_T^{miss} is that H_T^{miss} sums jets with $p_T > 30 \text{ GeV}$ and is blind to objects with $p_T < 30 \text{ GeV}$. Since we rely in our background estimation on jets with p_T in the range of $[15, 30] \text{ GeV}$, H_T^{miss} is more appropriate than E_T^{miss} since it does not introduce a bias in the data-driven background estimation methods. The two observables are highly correlated though, and describe similar physics.
- $N_{\text{jets}}(p_T \geq 30 \text{ GeV and } |\eta| < 2.4) \geq 1$. We are requiring at least one jet in the event, since such an ISR jet gives a boost to the produced neutralino, thus increasing the missing transverse energy and with it the sensitivity.
- $N_{\text{b-jets}}(p_T \geq 30 \text{ GeV and } |\eta| < 2.4) = 0$. We are vetoing any b-tagged jet. Our signal does not contain real b-tagged jets. This veto is efficient in rejecting background from $t\bar{t}$, in which the b quarks arise from a t quark decay.
- $\min \Delta\phi(H_T^{\text{miss}}, \text{jets}) > 0.4$. Since we are requiring an ISR jet in the event, we expect the H_T^{miss} to point in the opposite direction of the jet, or at least in an angle close to π . Events with multiple jets in the SM background such as arising from QCD will not exhibit such a feature. Therefore this cut reduces QCD background.
- veto events with isolated loose-ID lepton having $p_T \geq 30 \text{ GeV}$. Lepton can be either muon or electron. Our signal does not have high- p_T leptons, as we have seen in 7.5.
- $0.4 < m_{\ell\ell} < 12 \text{ GeV}$. Our signal resides in an invariant mass window with an edge at the mass difference between $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$. This is a relatively loose cut that is expected to further be tightened by the boosted decision tree.

We have already seen the object level selection in 7.7. For the sake of completeness we reiterate them here. The electrons in the analysis require to pass the following selection (see also 7.7.1):

- $5 \leq p_T \leq 15 \text{ GeV}$
- $|\eta| < 2.5$
- pass jet isolation

- loose ID

The muons in the analysis require to pass the following selection (see also 7.7.2):

- $2 \leq p_T \leq 15 \text{ GeV}$
- $|\eta| < 2.4$
- pass jet isolation
- medium ID

7.9.2 Category selection

The analysis has two main categories: dilepton and exclusive track category. The dilepton category has two leptons, while the exclusive track category has a single lepton and a track that has not been identified as a lepton. The dilepton category includes only one flavor, namely muons, while the exclusive track category has both electrons and muons flavored leptons. Selection for the dilepton category is listed in 7.9.2.1 and for the exclusive track category in 7.9.2.2.

7.9.2.1 Dilepton selection

In the dilepton category we have two reconstructed and identified muons. The following lists the selections:

- $N_\mu = 2$ opposite charge passing the muons selection.
- $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R(\mu_1, \mu_2) < 0.3$. This requirement makes this analysis orthogonal to the SOS analysis [1].
- event level BDT cut of $\text{BDT} > 0$. This is the main method of selecting signal events while rejecting the SM background. See 7.9.3 for details.
- $\Delta R(\mu_{1,2}, \text{leading jet}) > 0.4$. The leptons should not be inside the ISR jet.
- ω, ρ^0 and J/ψ invariant mass vetoes. $m_{\ell\ell} \notin [0.75, 0.81] \text{ GeV}$, $m_{\ell\ell} \notin [3, 3.2] \text{ GeV}$.

7.9.2.2 Exclusive track selection

In the exclusive track category we have one reconstructed and identified lepton (either an electron or a muon) and an exclusive track, i.e., a track that does not match an identified lepton. The track in the event that is picked to act as the misidentified lepton is the one with the maximum score of the track BDT 7.7.6. The following lists the selections:

- $N_\ell = 1$ lepton passing the muons selection.
- track picking BDT cut of $\text{BDT} > 0$. See 7.7.6.
- event level BDT cut of $\text{BDT} > 0$. This is the main method of selecting signal events while rejecting the SM background. See 7.9.3 for details.
- $\Delta R(\ell, \text{leading jet}) > 0.4$. The lepton should not be inside the ISR jet.

7.9.3 Boosted decision trees

In order to reject Standard Model (SM) background, pick signal events and define Signal Regions (SRs), in this analysis we employ multivariate method of type Boosted Decision Tree (BDT). For the dimuon category we train one BDT. For the exclusive track category we train a BDT for each lepton flavor and also for the two phases of the tracker detector, i.e., phase 0 and phase 1. That makes five BDTs overall.

All BDTs use the same structure of 120 trees with a maximum depth of 3, with the TMVA package [8]. The BDT training is performed with AdaBoost and GiniIndex separation. We are taking all other values as the defaults set by the TMVA package.

For the training we take tracks from a pool of our privately produced FASTSIM signal simulations which were listed in 7.6.2 for the signal, and the standard model background simulation listed in 7.6.1 for the background. For the exclusive track category, since we train for the different phases of the tracker detector, we use 2016 simulations of the signal and standard model background to represent phase 0, and the 2017 ones to represent phase 1. For the dimuon category we use only 2017 simulations to represent both phases with an added systematic uncertainty that results from this choice.

For the pool of our privately produced FASTSIM signal simulations, we are selecting the full range of simulated higgsino parameter μ (or the mass of $\tilde{\chi}_1^\pm$ in case of phase 1), but only the range of Δm we want to be most sensitive to. In phase 0, we select $\Delta m^0 \in [0.3, 4.3]$ GeV and $\mu \in [100 - 130]$ GeV. In phase 1 we select $\Delta m^\pm \in [0.3 - 4.6]$ GeV and $\mu \in [100 - 500]$ GeV. The baseline selection listed in 7.9.1 has been applied. A subset of the selections listed in 7.9.2.1 and 7.9.2.2 is used for the training:

- $N_\mu = 2(1)$ opposite charge passing the muons selection for the dimuon category (for the exclusive track category).
- $\Delta R(\ell, \text{leading jet}) > 0.4$.
- track picking BDT cut of BDT > 0 for the exclusive track category.

The training has been done without using MC weights, since we found that it gives the result with the least amount of over training that arise due to high weighted MC events. The training results in satisfactory performance since the kinematics of the low- p_T leptons is expected to look similar in all SM background processes. In the following sections, this fact needs to be taken into account when we look at the training input variables distributions. These distributions are plotted as they are seen by the training algorithm, i.e., without MC weights and with signal events taken from a pool of different parameter values as described above. Therefore, the ROC curves cannot be understood as a simple signal efficiency versus background rejection. Every BDT output working point results in different signal efficiency depending on the signal parameters values. As we will see later on, we do not use one BDT value with a simple cut and count, but rather we are binning the SRs in BDT output values. We therefore choose to plot the ROC curve with a default cut of 0.0 for the sake completeness. To fully estimate the power of the training, we need to look at the significance when each signal point has been properly weighted together with the SM background processes.

7.9.3.1 Dimuon category

The training samples for the dimuons category contain 4350 signal events and 21842 background events. The training samples are then tested against the test samples of equal size. The distributions of the testing samples overlay on the training samples, as well as the ROC curve, are seen in 7.35. No significant over training is observed.

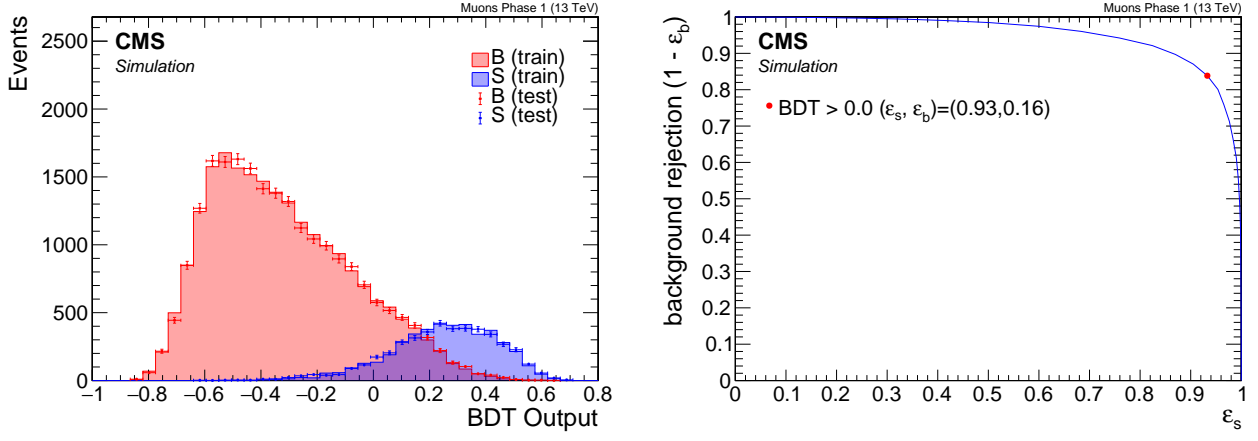


Figure 7.35: Dimuon BDT output (left) and ROC curve (right).

The training uses 18 different variable listed in 7.9.3.1 in decreasing order of importance ranking.

Table 7.9: Dimuon BDT input variables

Rank	Variable	Description
1	$m_{\ell\ell}$	invariant mass
2	$p_T(\ell_1)$	leading lepton p_T
3	H_T^{miss}	
4	H_T	
5	$\Delta R(\ell\ell)$	
6	$\min \Delta\phi(H_T^{\text{miss}}, \text{jets})$	
7	$p_T(\vec{\ell}_1 + \vec{\ell}_2)$	dilepton p_T
8	$p_T(\text{leading jet})$	
9	$p_T(\ell_2)$	subleading lepton p_T
10	$\eta(\ell_1)$	leading lepton η
11	$m_T(\ell_1)$	leading lepton transverse mass
12	$ \Delta\phi(\ell_2, \vec{H}_T^{\text{miss}}) $	
13	$ \Delta\phi(\ell_1, \vec{H}_T^{\text{miss}}) $	
14	$ \Delta\phi(\ell\ell) $	
15	N_{jets}	Number of jets
16	$\eta(\text{leading jet})$	
17	$ \Delta\eta(\ell\ell) $	
18	$m_{\tau\tau}$	collinear approximation of $m_{\tau\tau}$

Distributions of the input variables to the BDT training can be seen in 7.36. As mentioned before, the signal is taken from a pool of a range of model points, and events are not weighted to any luminosity or cross section in order to avoid over training.

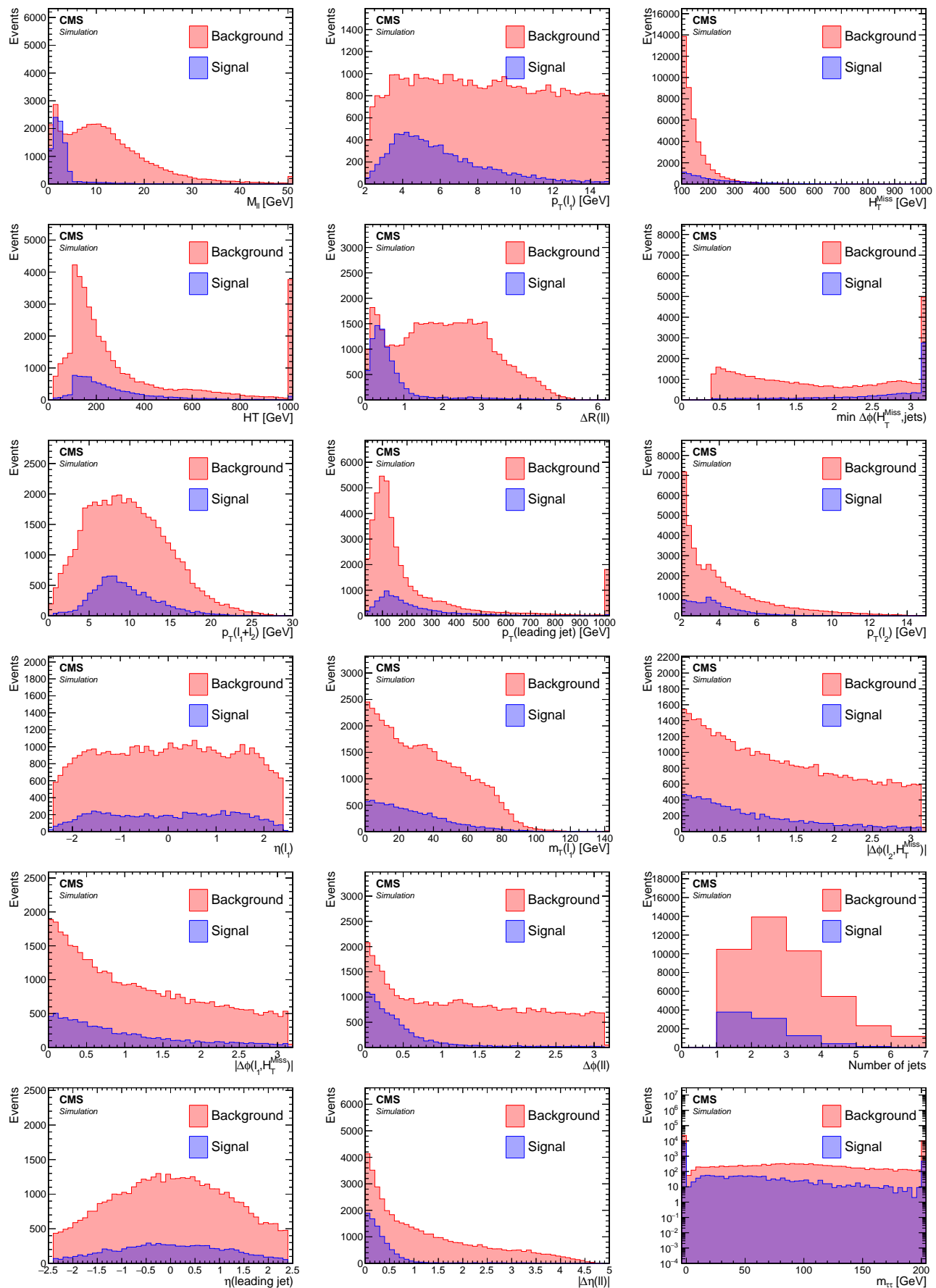


Figure 7.36: Dimuon BDT training input variables. Ordered by importance ranking.

7.9.3.2 Exclusive track category

The training samples in phase 0 for the exclusive category contain 7863 (1750) signal events and 55765 (29135) background events for muons (electrons) flavor. For phase 1, the exclusive category contain 5266 (1332) signal events and 51308 (31149) background events for muons (electrons) flavor. The training samples are then tested against the test samples of equal size. The distributions of the testing samples overlay on the training samples are seen in 7.37. No significant over training is observed. The ROC curves are seen in 7.38.

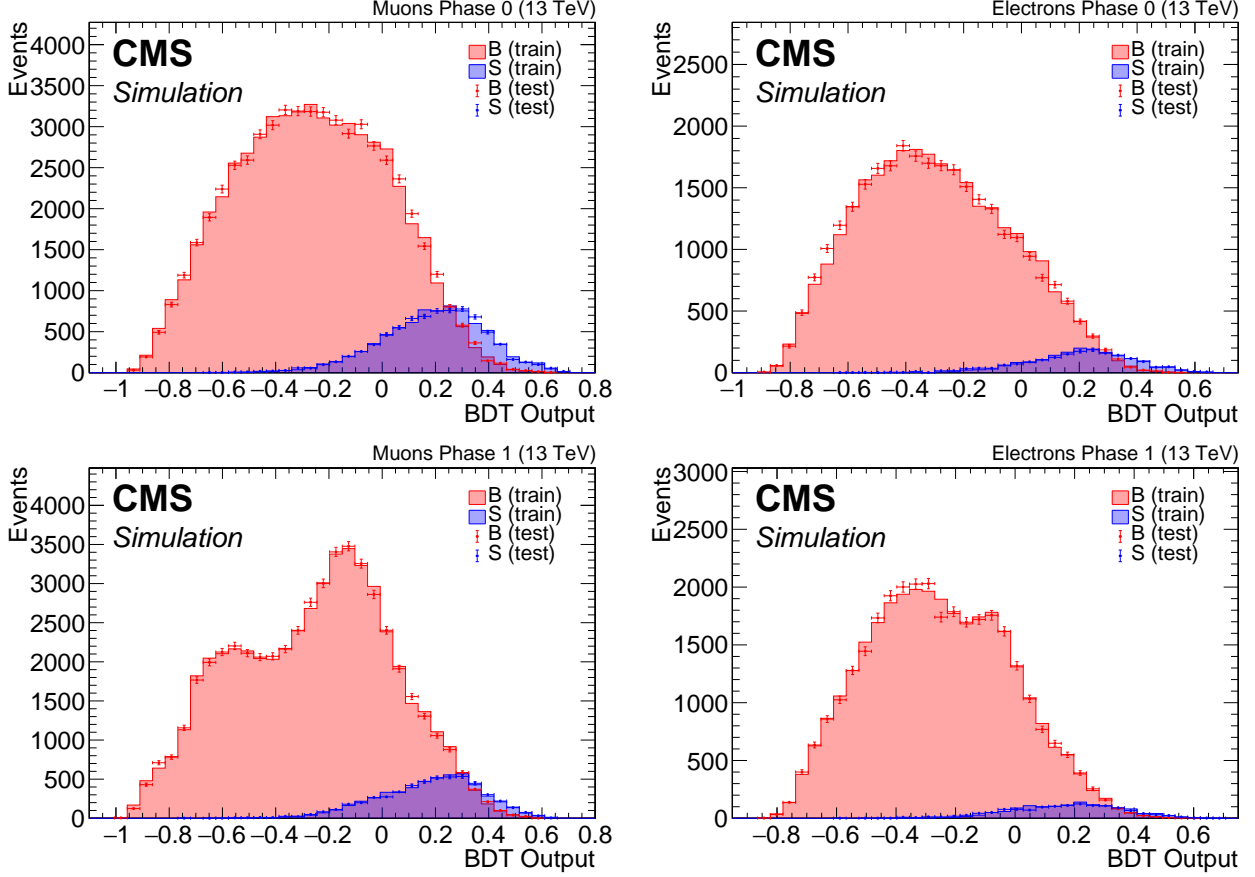


Figure 7.37: Exclusive track category BDT output in phase 0 (top) and phase 1 (bottom) for muons (left) and electrons (right)

The training uses 18 different variables listed in Table 7.9.3.2 in decreasing order of importance ranking. Since the ranking is slightly different in the four trainings, we choose here to list the order in the case of the muons of phase 1. We denote the fully identified lepton as ℓ and the non-identified lepton track as t .

Distributions of the input variables to the BDT training can be seen in 7.39. As mentioned before, the signal is taken from a pool of a range of model points, and events are not weighted to any luminosity or cross section in order to avoid over training. In the following sections we will look at fully weighted distributions in order to assess the performance of the training for different model points and to understand the different components of the standard model background and how to estimate it properly.

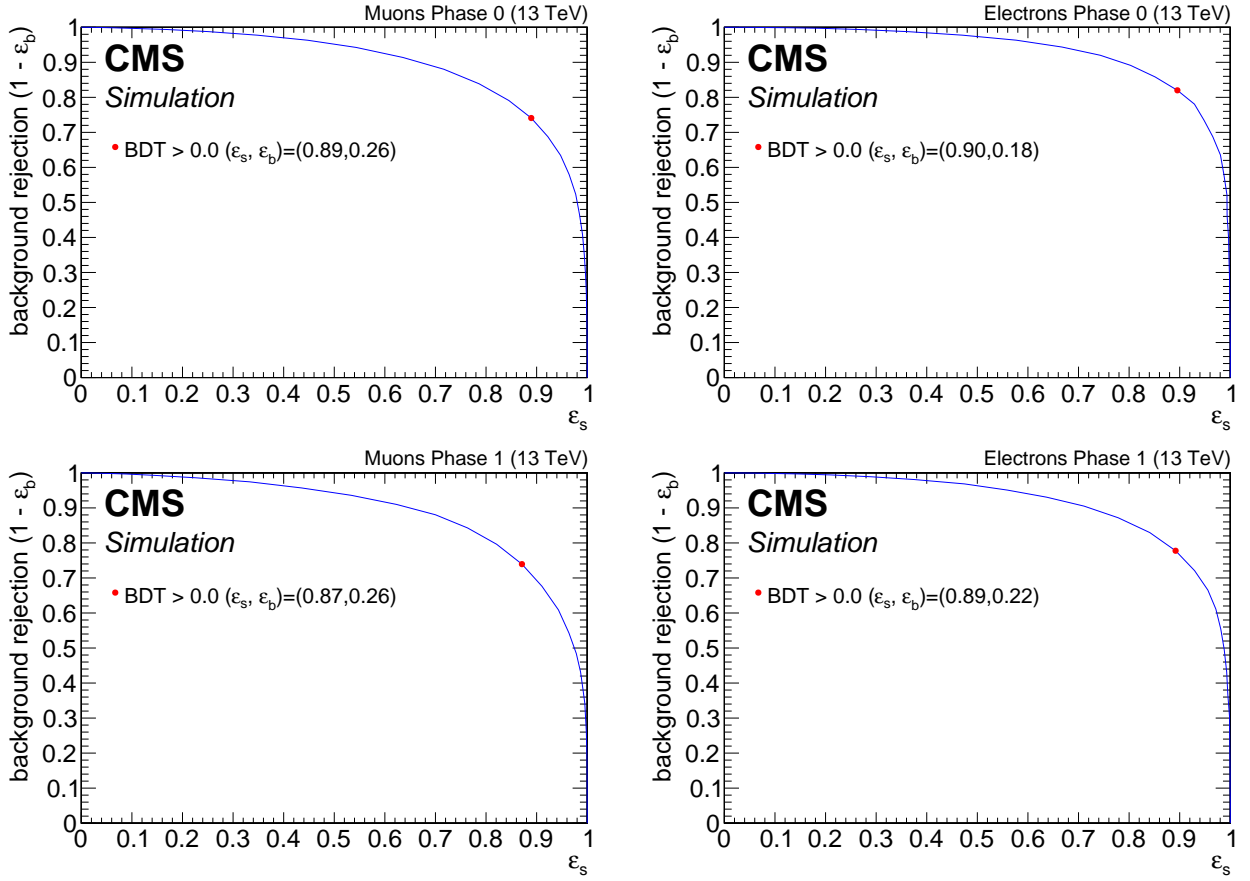


Figure 7.38: Exclusive track category ROC curves in phase 0 (top) and phase 1 (bottom) for muons (left) and electrons (right)

Table 7.10: Exclusive track BDT input variables

Rank	Variable	Description
1	$p_T(\ell)$	lepton p_T
2	H_T	
3	H_T^{miss}	
4	$\min \Delta\phi(H_T^{\text{miss}}, \text{jets})$	
5	$p_T(\text{leading jet})$	
6	N_{jets}	Number of jets
7	track BDT output	
8	$\eta(t)$	
9	$p_T(t)$	track p_T
10	$\eta(\text{leading jet})$	
11	$m_{\ell\ell}$	invariant mass
12	$\eta(\ell)$	
13	$m_T(\ell)$	lepton transverse mass
14	$\Delta R(\ell, t)$	
15	$\phi(\ell)$	
16	$\phi(t)$	
17	$ \Delta\phi(\ell, t) $	
18	$ \Delta\eta(\ell, t) $	

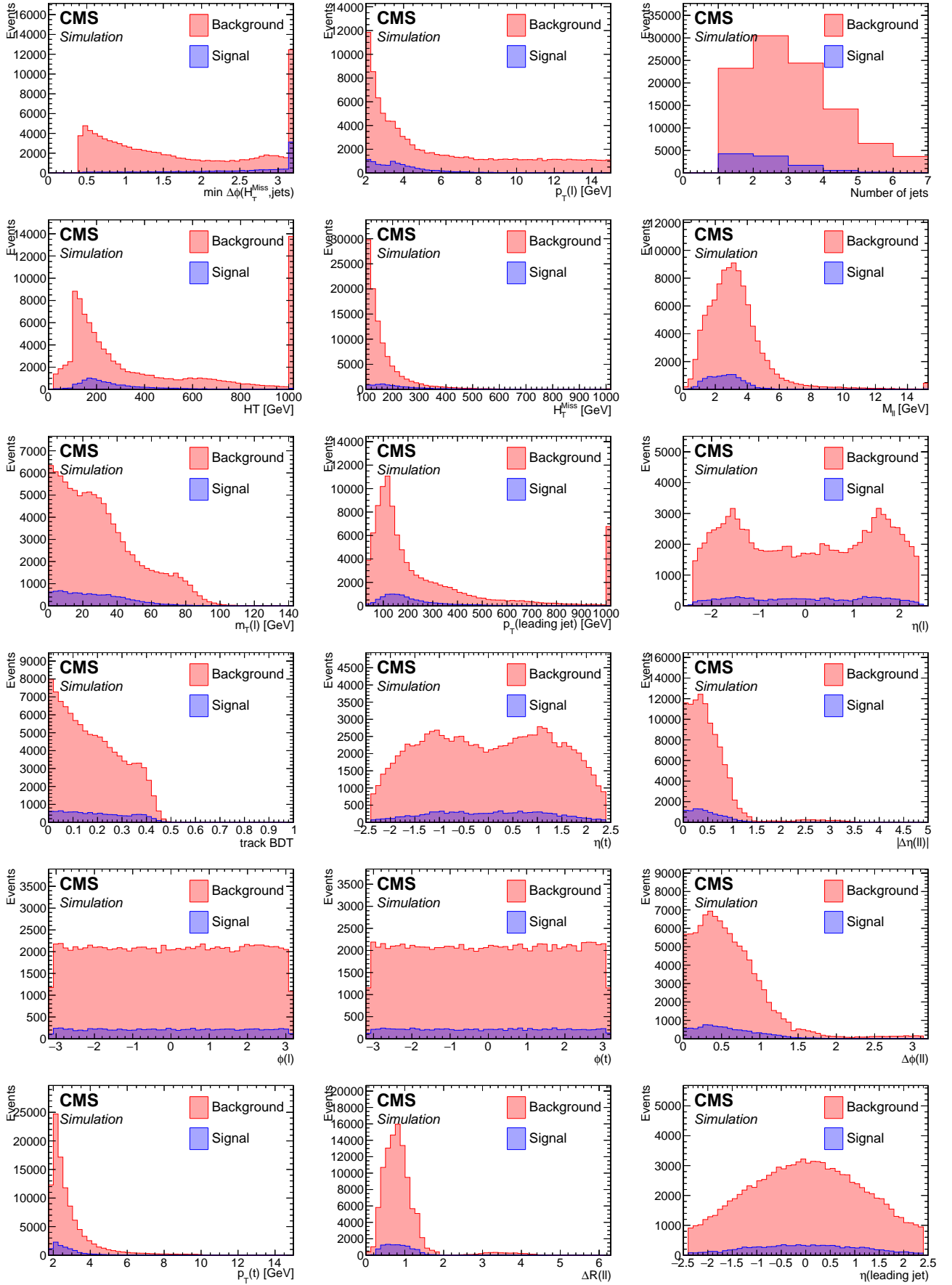


Figure 7.39: Exclusive track BDT training input variables. Ordered by importance ranking.

7.10 Characterization and estimation of the Standard Model backgrounds

In order to have a solid analysis strategy, the Standard Model (SM) background has to be thoroughly studied and estimated. It is important to fully understand the characteristics and composition of the SM background both in order to maximally reject it in the Signal Regions (SRs), as well as constructing a good method for estimating it. The characterization of the SM backgrounds is examined in 7.10.1, and estimation of the backgrounds is explored in 7.10.2.

7.10.1 Characterization of the Standard Model backgrounds

Signal Regions (SRs) are defined by a set of cuts or selections on observables composed of detector readings. Equivalently, multivariate techniques can be used in order to cut on a single aggregated output observable devised from a set of input variables, as was done in this analysis. No matter how a SR is defined, usually not only signal processes will be measured, but also processes that are not the intentions of the analysis. These processes are referred to as background to the analysis. Background processes can arise due to truly similar final state physics, or due to detector effects and mismeasurements. In an analysis such as the current one, an example for background in the dimuon category which arises from truly similar physics is Drell-Yan. In a Drell-Yan process, opposite-charge same-flavor dilepton pair are produced from an off shell Z^* or γ^* . An example of a background process that is due to mismeasurement is the production of a W in association with jets, which decays leptonically, and another lepton is due to either mismeasurement, i.e., a fake lepton, or as part of a hadronization procedure. A full list of the SM processes that are taken into account in this analysis with descriptions is given bellow. The processes are ordered according to their contribution in the SRs of the dimuon category.

- **W in association with jets.** In this SM process, we have a production of a W boson alongside jets, decaying leptonically into a lepton and a neutrino. It can be written schematically as $W + \text{jets} \rightarrow \ell \nu$. There are several reasons why this process is a background to this analysis. Since there's a neutrino produced in this process, it contributes to some real missing transverse momentum being produced. The presence of jets means that it is highly likely to pass the selection of at least one jet in the event. Energy mismeasurement of the jets can also contribute to the presence of fake missing transverse momentum, i.e., E_T^{miss} due to mismeasurements rather than real physics. Lastly, the very loose transverse momentum p_T range of the analysis muons means that it is also highly likely to pick an additional lepton in the event. In the exclusive track category, the extra track is mostly a fake lepton, i.e., a track that was misidentified by the BDT selection as the second lepton. In the dimuon category it is either a fake misidentified lepton, or a low- p_T lepton originating from a decay chain in a hadronization process.
- **Z in association with jets.** In this SM process, we have a production of a Z boson alongside jets, decaying into two neutrinos. It can be written schematically as $Z + \text{jets} \rightarrow \nu \nu$. The two neutrinos in this process contribute to true missing transverse momentum in the event. The two leptons will then come from either a decay of a meson produced in the hadronization procedure, or from unrelated decays. Another highly likely scenario is for one of the leptons to be fake, just like in the $W + \text{jets} \rightarrow \ell \nu$ case.
- **Drell-Yan process.** It takes place when a quark of one proton and an antiquark of another proton annihilate, creating a virtual photon or Z^* boson which then decays into a pair of oppositely-charged leptons. When two electrons are produced via $Z \rightarrow e^+ e^-$ or two

muons via $Z \rightarrow \mu^+ \mu^-$, true missing transverse momentum is not part of the production. Therefore, a relatively high E_T^{miss} cut such as in this analysis is successful in suppressing these types of backgrounds. However, in the production of two taus via $Z/\gamma^* \rightarrow \tau^- \tau^+$, each tau can then decay into a muon alongside two neutrinos, i.e., $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$, producing real missing transverse momentum in the event alongside two real leptons, which become a background to this analysis.

- **Ditop in association with jets.** When two top quarks are produced, $t\bar{t}$, most of the time they will decay into a W and a b quark. That happens close to 100% of the time. Therefore, $t\bar{t}$ becomes a background when each top quark decays into a b quark and a leptonically decaying W boson. The W produces neutrinos in the decay alongside the leptons, contributing real missing transverse momentum. Vetoing b-tagged jets, as listed in 7.9.1, is successful in suppressing most of this background, but not all.
- **Numerous bosons production.** In the plots of the following section, we differentiate between diboson production, VV, and higher order production such as three bosons, which are referred to collectively as *rare*. The ways in which they can be selected in the SRs are similar to the single boson case, only that the higher multiplicity cases have much lower production cross sections, and therefore are almost negligible in this analysis.
- **QCD production.** Quantum Chromodynamics (QCD) comprises events arising from the production and radiation of quarks and gluons followed by their hadronization and showering into highly columnar sprays of particles known as jets. QCD events contain no real E_T^{miss} . Any E_T^{miss} present in a QCD event is due to mismeasurements. The relatively high E_T^{miss} cut in this analysis in combination with requiring $\min \Delta\phi \left(H_T^{\text{miss}}, \text{jets} \right) > 0.4$, reduce almost all QCD background, and is therefore a very negligible background in this analysis. That is the reason why it can be safely disregarded in most contexts and safely assumed that any remaining events will be successfully be estimated using the jetty-background data-driven background estimation method in 7.10.2.1.
- **Resonances.** **FiXme Note: Add description here.**

The optimal way to gather an understanding of the proportion each background process has in the Signal Regions (SRs), is to look at correctly weighted MC simulation, where each process is weighted to its corresponding production cross section, and the overall background is weighted to match the era luminosity. It is useful to look at the final SRs defined and discussed in Section 7.12, but also at some of the important inputs to the BDT which were discussed in 7.9.3. In the following sections, it is useful to note that the SRs are all the bins with a BDT value greater than zero. Each of the following sections presents the figures relevant to the relevant analysis category, beginning with the BDT output distributions and SRs bins, following by important inputs to that BDT. Although input distributions have been seen already in section 7.9.3, here the background processes and individual signal models are properly weighted. The data-taking year of 2017 was chosen as a representative here.

7.10.1.1 Dimuon category

The dimuon 2017 simulation BDT output, weighted to 2017 luminosity, both in log and linear scales, can be seen in Figure 7.40. The SRs are the six bins with BDT output greater than zero in increasing sensitivity. As can be seen in the figure, the largest backgrounds are $t\bar{t}$, $Z + \text{jets} \rightarrow \nu\nu$ and $W + \text{jets} \rightarrow \ell\nu$. There is also a small presence of Drell-Yan processes, that are mainly due to $Z/\gamma^* \rightarrow \tau^-\tau^+$. The top ten input observables to the BDT, ranked by importance for the training, can be seen in Figure 7.42.

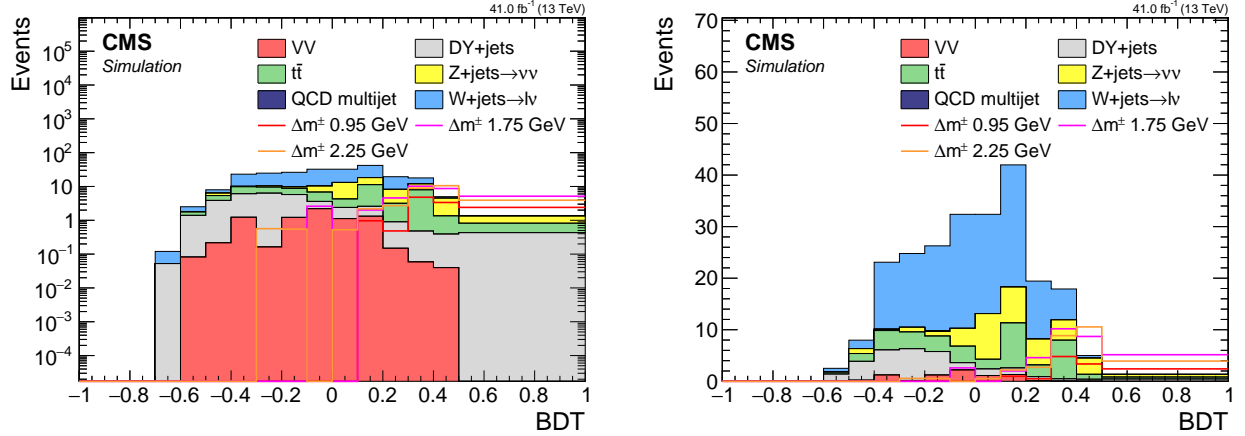


Figure 7.40: Dimuon 2017 simulation BDT output in log scale (left) and linear scale (right).

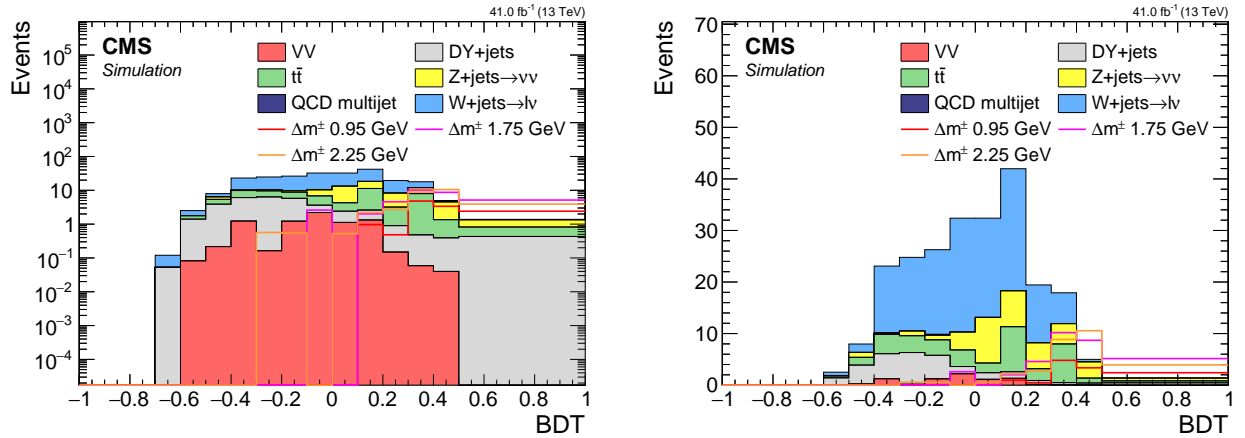


Figure 7.41: Dimuon 2017 simulation BDT output in log scale (left) and linear scale (right). Composition split into...

FiXme Note: Add jetty/tautau simulation plot

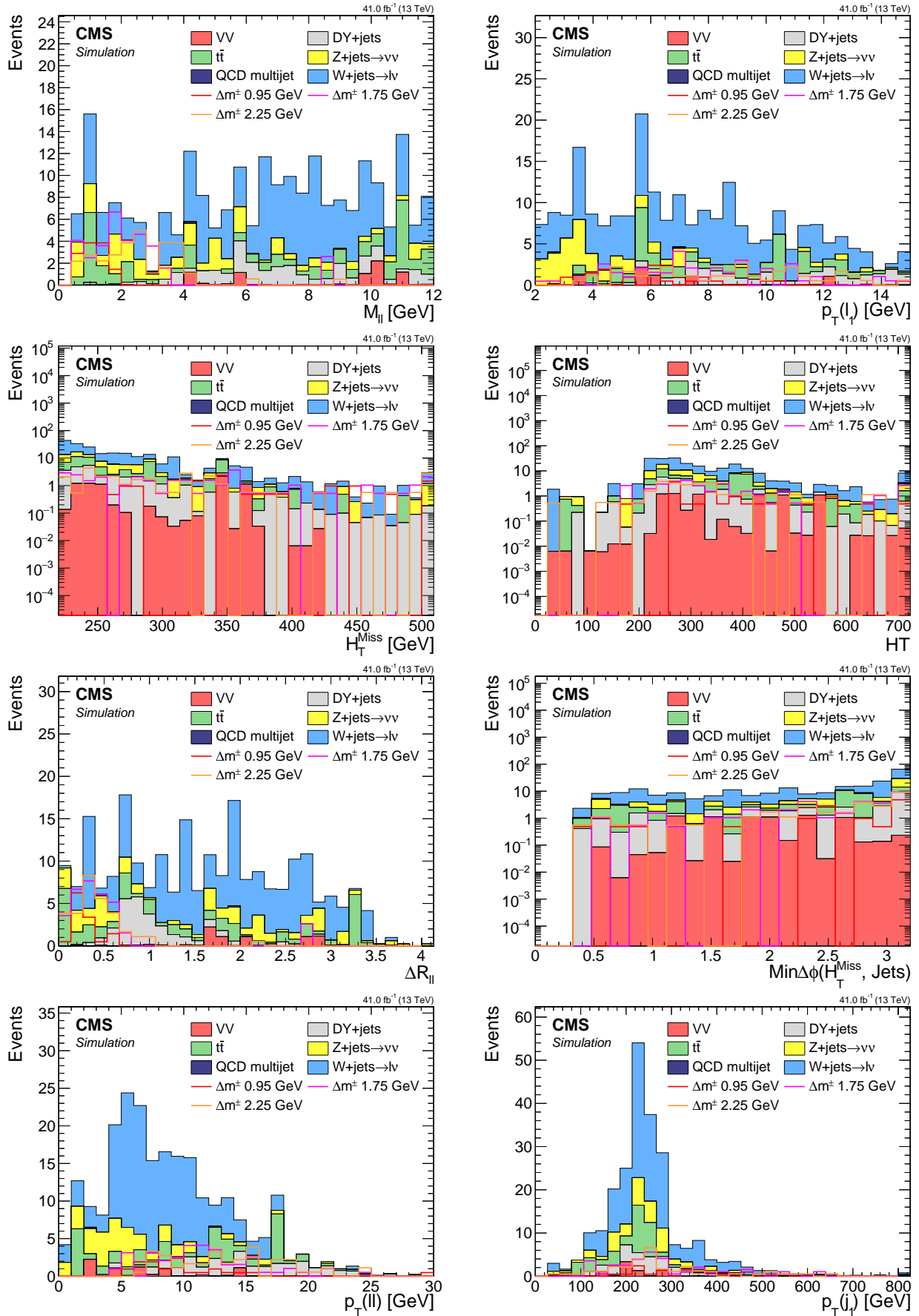


Figure 7.42: Dimuon 2017 simulation BDT inputs for the top 10 ranked observables.

7.10.1.2 Exclusive track category

As described before, there are four BDTs in the exclusive track category, for each of the two lepton flavors there are two track phases.

The muon flavor in phase 1 is chosen to represent background composition in the exclusive track category in order to avoid repetition. Figure 7.44 shows the top eight input observables to the BDT, ranked by importance for the training. It is weighted to 2017 luminosity and uses 2017 simulation. A few signal points are also plotted in order to point to the signal-like regions of phase space.

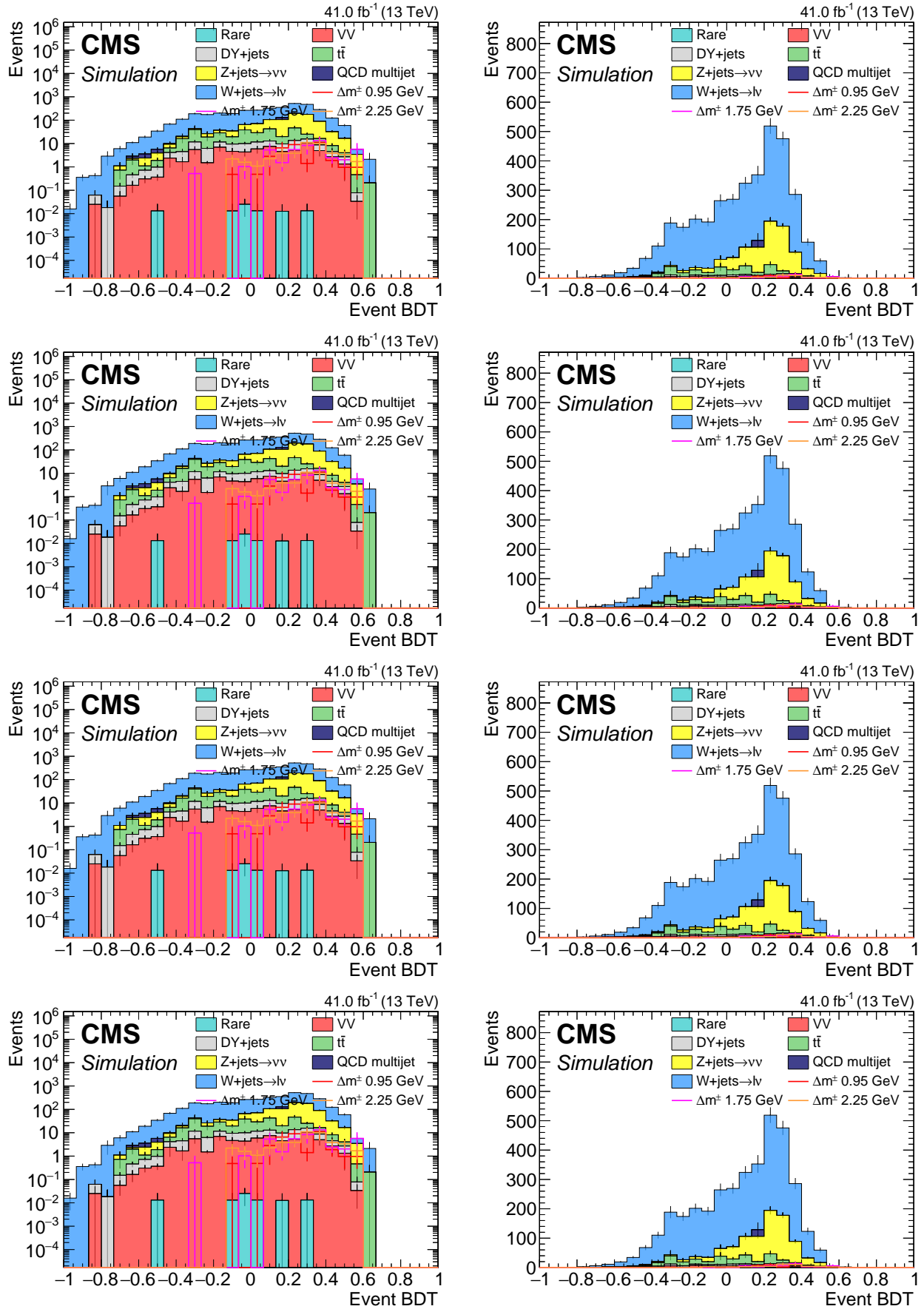


Figure 7.43: Dimuon 2017 simulation BDT output in log scale (left) and linear scale (right).

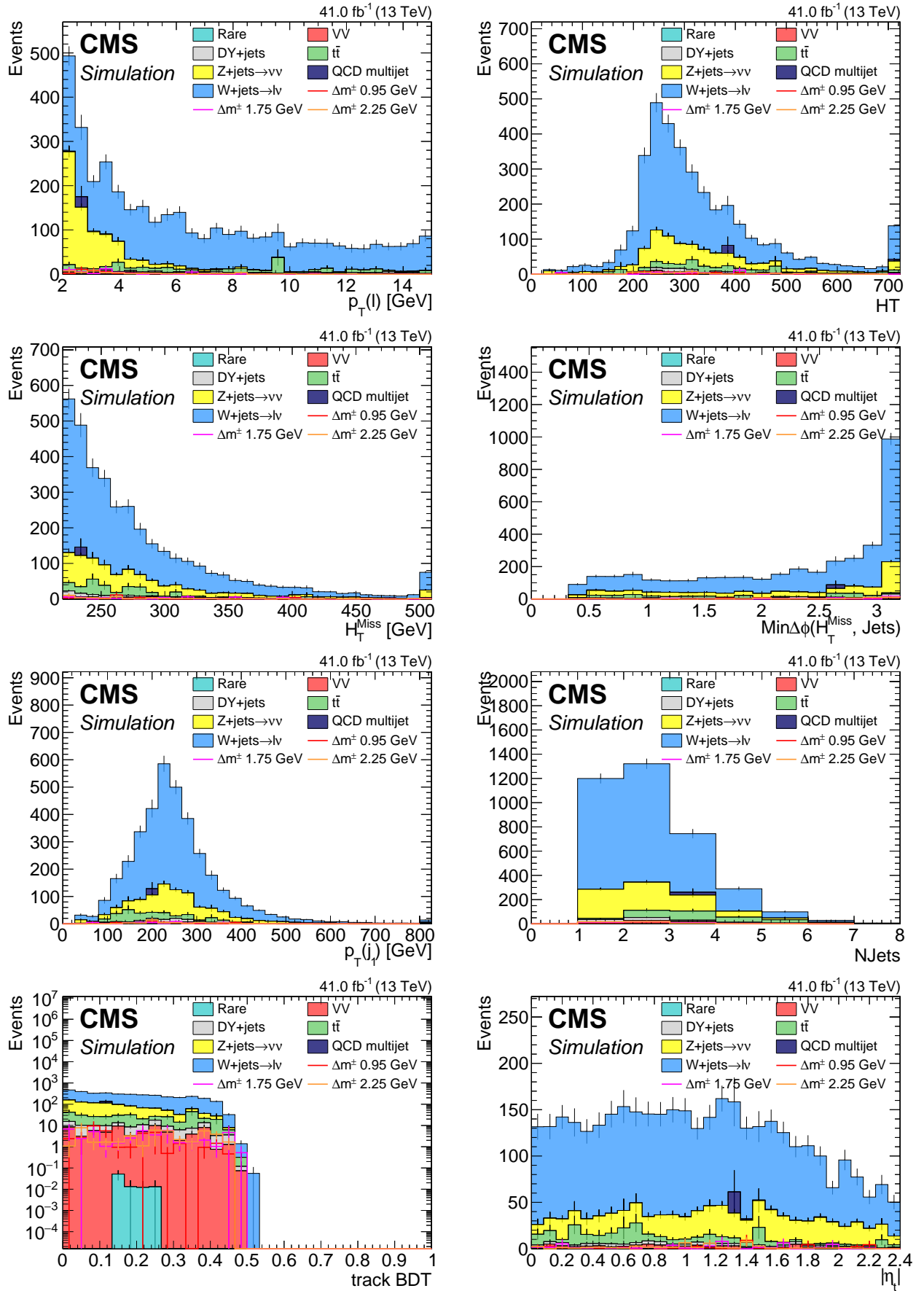


Figure 7.44: Exclusive track plus muon 2017 simulation BDT inputs for the top 8 ranked observables.

7.10.2 Estimation of the Standard Model backgrounds

In analyses such as the one described in this thesis, in which new physics, or physics beyond the Standard Model is searched for, one needs to predict event counts for the signal and for the Standard Model background. These counts are denoted s for the signal events count and b for the background events count. From these counts and their uncertainties, a *significance* is computed. The question that arises is how to predict such counts. A widely used method is Monte Carlo (MC) simulation, sometimes simply referred to as *simulation*. Prediction using simulation is pretty straight forward. First, many events are simulated using the SM technique, taking into account the detector's geometry (by such algorithms such as FASTSIM or FULLSIM). They are then weighted to account for production cross sections and luminosity. Further correction factors and weights can apply to account for measurements errors, discrepancies between data and simulation, non-ideal algorithms, misidentifications, normalization corrections computed from data in a dedicated Control Region (CR), and so forth. The nominal prediction for the events count becomes then the fully-weighted simulation count in the relevant Signal Region (SR) with an uncertainty that is taken by the statistical error on the simulation count in that bin, with an optional systematic uncertainty that has been determined by other means. The signal count s has been determined in this way using FASTSIM. As is described in the following, only a small fraction of the background count b has been estimated using simulation. The majority of the background processes have been estimated using *data-driven methods*.

Using simulation to estimate the Standard Model background has limitations and disadvantages. These are specific to each analysis, in which the background processes composition change. The main limitation of simulation is that it is imperfect. Simulation never precisely simulates real data, and that is due to several factors. Theoretical uncertainties, such as uncertainties on cross sections or branching fractions, can lead to wrong production rate or normalization. That's why simulation is often reweighted using a normalization factor derived in a dedicated CR. Another challenging limitation of simulation is misrepresentation of the detector's delicate geometry and response, as well as real time data-taking challenges and faults which need to be taken into account in order for the simulation to accurately represent the events that would have been recorded by the detector. Some objects and phase-spaces are more prone to discrepancies than others.

In this analysis, a great deal of the challenge arises from the very *soft* nature of the leptons, i.e., their low transverse momentum p_T , as well as their low invariant mass of the order of a few GeV. Sources of backgrounds of such events in the standard model include low- p_T resonances that are produced in hadronization processes, and events where one of the leptons (or exclusive track) is misidentified as one of the signal leptons. These leptons or tracks are often in some vicinity of jets in the event. In this analysis, there are two strategies to estimating this sort of background depending on the whether we have two identified leptons, such as in the dimuon category, or one, such as in the exclusive track category. The dimuon *jetty background* estimation is described in Section 7.10.2.1, while the exclusive track background estimation is described in 7.10.2.3. Lastly, as described earlier, small part of the background is not produced nearby jets at all, namely $Z/\gamma^* \rightarrow \tau^- \tau^+$. That background estimation method is described in Section 7.10.2.2.

7.10.2.1 Non-isolated jetty background estimation

In the description of the object selection in Section 7.7, it has been demonstrated that the leptons in the signal are isolated. The isolation criterion that is used in this analysis has been described in 7.7.7, and is referred to as *jet-isolation*. It has been mentioned there that the customized isolation is also key part in the background estimation, and that is now described in this section. This background estimation method is relevant to the dimuon category. It is a *data-driven* background estimation method, meaning that we use data, rather than simulation, to estimate this background. The name *non-isolated jetty background* refer to background in which one or both of the leptons generally are produced in association with jets, and typically are close to one. Most of such leptons are rejected by the jet-isolation criteria, but some do manage to pass the isolation if they happen to have been produced far enough from a jet.

Our signal regions are defined in detail in Section 7.12, but here it is enough to note they are bins in the BDT output with a score greater than zero. The strategy of this method is to find a Control Region (CR) that produces a BDT output distribution that is shape-consistent with the original BDT distribution. That is achievable when a CR can be identified such that it has almost the same features as the main region, ideally only with rate difference. Then, to get normalization correctly, one would define a normalization control region to correct for the different production rate.

Interim note on terminology should be made for clarity. The SRs are defined by taking BDT output greater than zero, and therefore, by definition, the region with less than zero becomes a CR. Furthermore, a new region using isolation is being introduced in this section. Since the orthogonal region only varies from the main region by isolation, it is referred to as the *isolation sideband*. The main region that was used to define the SRs using the BDT output is referred to as the *main band*. The SRs are then regions in the *main band* with BDT output greater than zero. The events in the *isolation sideband* are used to predict the jetty-background in the *main band*. The *normalization region* is taken to be in the CR with $\text{BDT} < 0$, and can also be referred to as *BDT sideband* or more elaborately as *BDT normalization sideband*. Of course though, a *sideband* is still a kind of a CR.

Since the goal of this estimation method is to estimate the jetty-background, an isolation sideband is defined by requiring either one or two of the leptons to fail the jet-isolation criterion. That has also been defined in step 5 in Section 7.7.7. That means, that the lepton is distanced less than 0.6 in ΔR from a lepton-corrected jet. The jet that caused the lepton to fail the jet-isolation is required to have an original transverse momentum, i.e., transverse momentum before the subtractions, satisfying $15 < p_T < 30$ GeV. The reason why the p_T is limited to this region is because 15 GeV is the upper bound on the leptons, and 30 GeV is the lower bound on the objects being summed in the calculation of H_T^{miss} and the number of jets in the event. If the upper bound of 30 GeV had not have been set, a bias in the isolation sideband would have been introduced, because requiring a lepton to fail jet-isolation, would also potentially mean requiring an extra jet with $p_T > 30$ GeV, a requirement that didn't exist in the main band. The distributions of number of jets, as well as of H_T^{miss} , are blind to jet with $15 < p_T < 30$ GeV, which makes it safe to potentially have extra jets in that region. The BDT training will be blind to those jets, and the BDT shape in the sideband should not be affected by those, thus achieving the goal of having consistent shapes between the main band and the sideband.

The main assumption that goes into using the isolation sideband is that in the background, the leptons are not isolated, but created in association with jets. Most will be produced inside jets, with a smoothly falling distribution outwards from a jet. By setting a cone of size 0.6 around jets and picking leptons inside those cones, events are picked up which have similar behavior to events where the leptons are outside of those cones. The rate of production of leptons inside jets will be much higher than those outside of the jets, but since their behavior

is similar kinematically, by fixing the normalization, their distributions will agree with the ones in the main band. The normalization factor is calculated by taking the ratio between the main band and the isolation sideband in the CR of $\text{BDT} < 0$. The event count in the sideband is multiplied by the normalization factor to comprise the prediction. The prediction in the SR then becomes:

$$N_{\text{predicted}}^{\text{SR}} = \frac{N_{\text{main band}}^{\text{CR}}}{N_{\text{sideband}}^{\text{CR}}} \cdot N_{\text{sideband}}^{\text{SR}}. \quad (7.10)$$

The assumption that the isolation sideband, i.e., events with at least one of the leptons failing the jet isolation criterion, correctly predicts the shape of the main band in the signal region, is tested by performing a shape comparison in simulation. Such shape comparison, or *closure test*, is performed by plotting the main band and the jet isolation sideband simultaneously, and checking that they statistically agree in terms of shape. A normalization factor is computed in order to correctly normalize the isolation sideband. In the current section, treatment of phase 1 (2017-2018), which in simulation is represented by 2017 MC is presented. Special care has been taken in the case of phase 0 (2016 data taking year) in Section 7.13.

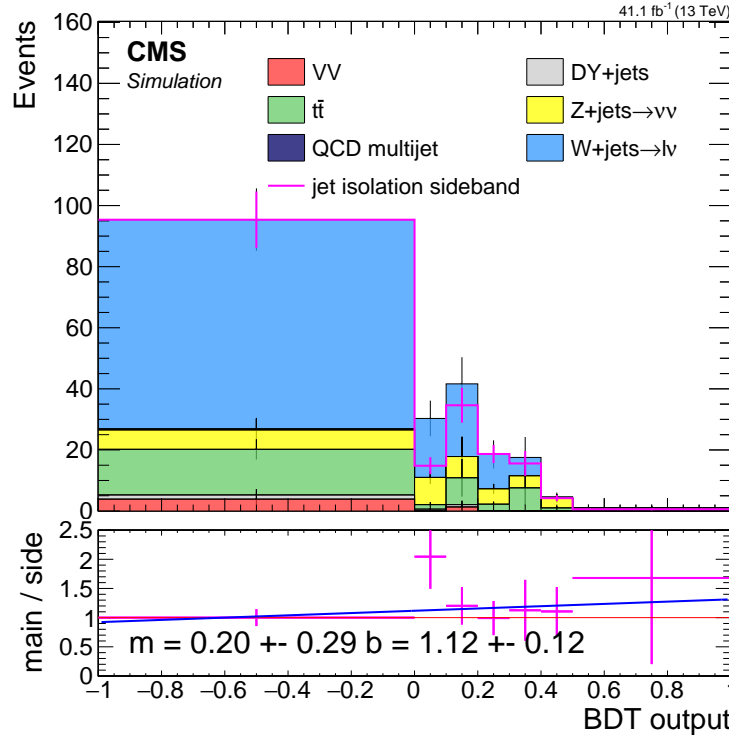


Figure 7.45: Dimuon 2017 jetty background closure plot. The stack represents simulation in the main isolation band after $Z/\gamma^* \rightarrow \tau^- \tau^+$ has been removed, while the pink line represents simulation in the isolation sideband. The isolation sideband is normalized to match the isolation in the CR of $\text{BDT} < 0$. The ratio panel shows the ratio between the isolation main band and sideband. A line fit of the ratio is performed and the parameters of the slope m and interception point b with their respective errors are stamped.

The jetty background closure plot is shown in Figure 7.45. The closure plot tests the assumption that the background in the isolation main band can be predicted by the isolation sideband. Therefore, after proper normalization that is achieved by dividing the event count in the the main band with the event count in the sideband, which in this case is 0.59, a flat ratio distribution is to be expected in the SRs. One can observe that indeed such a flat ratio distribution has been achieved by noting that most bins are statistically consistent with unity.

To check for any trend, a line fit has been performed for the ratio panel. Taking into account the errors, up to 1σ , the slope m is consistent with 0 and the interception b is consistent with 1, making the line consistent with a flat line intercepting 1. It is to be concluded therefore, that the closure plot confirms the shape assumption, and no trend that needs to be taken into account is to be seen.

7.10.2.2 Ditau Drell-Yann background estimation

It has been seen in Section 7.10.1.1 that the majority of the standard model background is composed of processes that produce leptons in association with jets, and therefore can be estimated with the jetty-background data-driven estimation method described in Section 7.10.2.1. However, as can be seen in Figure 7.41, a small amount of background arising from $Z/\gamma^* \rightarrow \tau^- \tau^+$ is also present, and since the leptons due to the leptonic decay $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$ are isolated, it requires an alternative background estimation method.

The $Z/\gamma^* \rightarrow \tau^- \tau^+$ background in this category is estimated using MC simulation. The event counts is being weighted according to a data correction normalization factor calculated in a dedicated CR. In order to do so, a CR rich and relatively pure in $Z/\gamma^* \rightarrow \tau^- \tau^+$ background must be found. In order to achieve this task, the observable $m_{\tau\tau}$ must be introduced. The background in question is due to a Drell-Yann process where a Z is decaying into two tau leptons, which in turn decay into muons via $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$. If the taus could be reconstructed, they could be used to calculate the invariant mass of the ditau event, $m_{\tau\tau}$, which is expected to peak at around the Z mass. The Z resonance could then be used as the desired CR rich in ditau background. However, leptonic taus are not directly reconstructed at CMS, which makes it necessary to find an alternative approach for calculating said invariant mass.

A widely used method for the reconstruction of the invariant mass $m_{\tau\tau}$ is the so-called *collinear approximation*. First described in [12], it has been used widely in ATLAS [13] and CMS [14]. In the collinear approximation, one assumes that the tau pair produced from Z/γ^* are highly energetic, so that their leptonic decays are collinear and the source of missing transverse momentum is due to the neutrinos only. If both the τ -leptons are sufficiently boosted, the neutrinos from each τ decay are collinear with the visible lepton momentum. One can then use the visible daughter-lepton momentum together with \vec{p}_T^{miss} to reconstruct the τ -lepton pair and calculate the would-be $m_{\tau\tau}$ invariant mass. Depending slightly on the mathematical details of the approximation, one can arrive at a strictly positive distribution for $m_{\tau\tau}$, as was done for example in [15], or one that also has negative values, as was done in [16, 17]. Since the negative values arising from the second variation of the approximation are due to events where \vec{p}_T^{miss} points away from one of the leptons, and therefore not arising from underlying boosted ditau events, it is useful to reject negative values in order to purify the CR by rejecting those events. The collinear approximation does not work when the taus are back-to-back. However, since in this analysis an ISR jet is required together with a high threshold of missing transverse momentum, it is expected to work well on the remaining boosted events. The signal, as well as other standard model processes are expected to have a flat distribution in $m_{\tau\tau}$, while events arising due to $Z/\gamma^* \rightarrow \tau^- \tau^+$ are expected to peak around the Z boson mass.

To put the assumptions into practice, an explicit definition of $m_{\tau\tau}$ is now laid out. The definition of the invariant mass is:

$$m_{\tau\tau}^2 = (p_{\tau_1} + p_{\tau_2})^2 \quad (7.11)$$

Assuming that the τ -pair is boosted, and the fully-leptonic decaying products is fully collinear to the τ -leptons, it follows that the transverse momentum of each neutrinos-pair is

proportional to the corresponding i τ_i 's transverse momentum by a scale-factor ξ_i by

$$\vec{p}_T^{\nu_i} = \xi_i \vec{p}_T^{\tau_i} \quad (7.12)$$

Since by assumption, all of the missing transverse momentum is due to the neutrinos, it follows that:

$$\vec{p}_T^{\text{miss}} = \xi_1 \vec{p}_T^{\tau_1} + \xi_2 \vec{p}_T^{\tau_2} \quad (7.13)$$

One solves the above two equations 7.13 for the two parameters ξ_1 and ξ_2 for each event. The solution becomes:

$$\begin{aligned} \xi_1 &= \frac{\vec{p}_T^{\text{miss}} \cdot \vec{p}_y^{\ell_2} - \vec{p}_T^{\text{miss}} \cdot \vec{p}_x^{\ell_2}}{\vec{p}_x^{\ell_1} \cdot \vec{p}_y^{\ell_2} - \vec{p}_x^{\ell_2} \cdot \vec{p}_y^{\ell_1}}, \\ \xi_2 &= \frac{\vec{p}_T^{\text{miss}} \cdot \vec{p}_x^{\ell_1} - \vec{p}_T^{\text{miss}} \cdot \vec{p}_y^{\ell_1}}{\vec{p}_x^{\ell_1} \cdot \vec{p}_y^{\ell_2} - \vec{p}_x^{\ell_2} \cdot \vec{p}_y^{\ell_1}}. \end{aligned} \quad (7.14)$$

Equation 7.11 is expended with the assumption that the τ 's are boosted and that the four-momenta of the taus is $p_{\tau_i} = (1 + \xi_i)p_{\ell_i}$:

$$\begin{aligned} m_{\tau\tau}^2 &= (p_{\tau_1} + p_{\tau_2})^2 \\ &= ((1 + \xi_1)p_{\ell_1} + (1 + \xi_2)p_{\ell_2})^2 \\ &= 2m_\tau^2 + 2(1 + \xi_1)(1 + \xi_2)p_{\ell_1} \cdot p_{\ell_2} \\ &\approx 2(1 + \xi_1)(1 + \xi_2)p_{\ell_1} \cdot p_{\ell_2}. \end{aligned} \quad (7.15)$$

This can be negative is $\xi_i < -1$. This can happen if the missing transverse momentum vector nearly opposite to a lepton's \vec{p}_T and also $p_T^{\text{miss}} > p_T^\ell$. This can easily happen in other background processes, such as $WW + \text{jets}$, when a neutrino and a lepton (possibly coming from different decay legs) are nearly back-to-back. Therefore, the final definition of $m_{\tau\tau}$ is

$$m_{\tau\tau} = \text{sign}(m_{\tau\tau}^2) \sqrt{|m_{\tau\tau}^2|}. \quad (7.16)$$

The CR to be used for the normalization should potentially have low signal contamination. It can be seen in Figure 7.40 that the region of $\text{BDT} < 0$ has hardly any signal contamination, and therefore is used for building the $\tau\tau$ CR. Figure 7.46 shows distributions of the $m_{\tau\tau}$ for the $\tau\tau$ simulation in red, and the rest of the standard model backgrounds in the stack. The two tracker phases are seen side by side. There is a clear peak in the $\tau\tau$ background at around the Z boson's mass. A window around the Z boson's mass is chosen of $[40, 130]$ GeV in order to achieve high purity of 73% in both phases. The contamination in data is removed by first predicting the jetty background count using the data-driven method described in Section 7.10.2.1, and subtracting those counts from the data counts in the $\tau\tau$ dedicated CR. Then, a data divided by simulation normalization factor is computed in said $m_{\tau\tau}$ window to get 1.2 ± 0.46 (0.29 ± 0.26) which has relative error of 38% (90%) for phase 0 (phase 1).

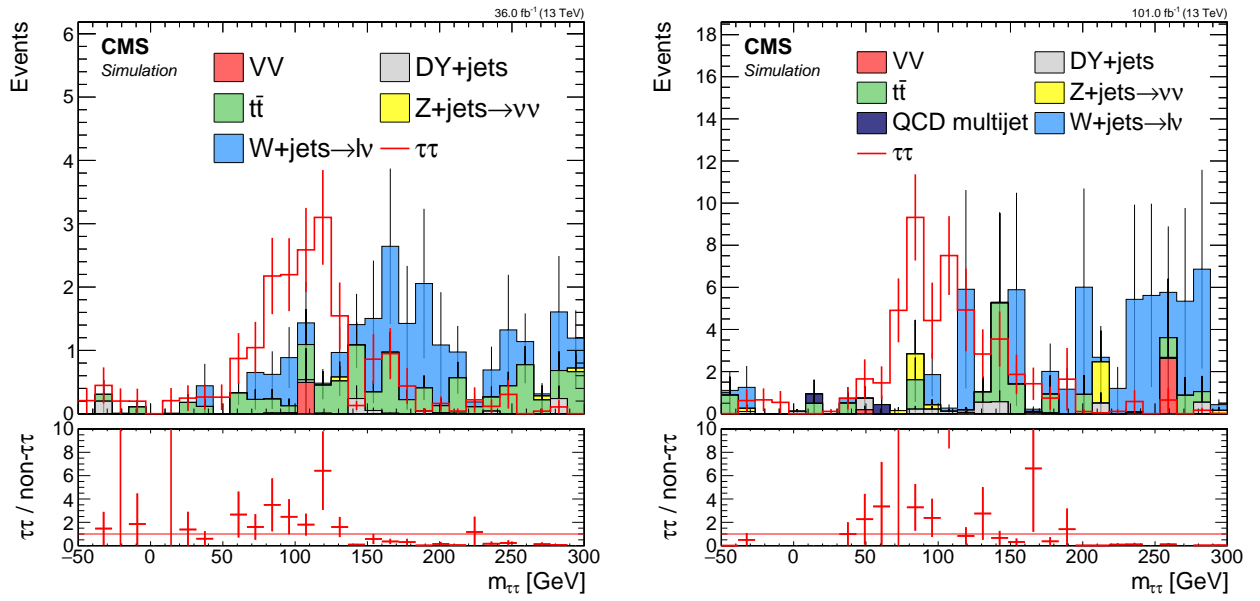


Figure 7.46: Ditau invariant mass $m_{\tau\tau}$ distributions for phase 0 2016 simulation (left) and phase 1 2017 simulation weighted to luminosity of 2017-2018 data taking period (right). The red line corresponds to $\tau\tau$ simulation, and the stack represents the rest of the standard model background simulation. No overflow bins are plotted in order to clearly show the resonance peak.

7.10.2.3 Exclusive track background estimation

The exclusive track category is consisted of four different BDTs, one for each lepton flavor and for each of those one for each phase. However, the background estimation method is the same for all of them, and it is a data-driven one.

The exclusive track category required one fully identified lepton according to the selection listed in Sections 7.7.1 and 7.7.2, and one track by a procedure described fully in Section 7.7.6. The track that is selected to be paired with the single identified lepton is selected as the track in the event with the highest BDT score among all tracks using a BDT that was trained on signal events in order to reject in-signal background and potentially pick up the track that corresponds to the non-identified second lepton in the signal event. In the background, the same selection procedure applies. However, unlike in the signal, it cannot be expected that the pairing of the selected track has a special relationship to the identified lepton other than passing some selections on the phase space. The chance of selecting a track in the background, that in combination with the identified lepton form an opposite-sign same-flavor pair that is a result of a resonance decay, is vanishingly small. It is highly likely that the track corresponds either to a lepton produced independently of the identified lepton in the event, or that it doesn't correspond to a lepton at all. Such a track is referred to as a *fake* track, and almost all tracks in the event are fakes.

To devise a successful data-driven background estimation for the exclusive track category we take advantage of the fact that the tracks in the background are independent of the identified lepton in the event. The selection process considered only tracks with opposite charge to the identified lepton, but if the track is independent of the lepton, events with a track of the same charge will be picked at the same rate as opposite charge ones. Not only is it expected to be selected at the same rate, but the shape of the BDT output will also be blind to this choice, making it an excellent proxy to the opposite charge background.

A CR is defined by selecting same-charge lepton-track pair rather than opposite-charge as in the SR. It is expected that the same charge CR is consistent also by the normalization to the SR since it is expected that these events are selected at the same rate of each other. Despite of this, the normalization is fixed by calculating a normalization factor as the ratio between the opposite-charge to same-charge event count in a dedicated normalization sideband CR satisfying $\text{BDT} < 0$, and applying it to the same-charge event count in the SRs satisfying $\text{BDT} > 0$. In order to test the independence assumption and to demonstrate the correct shape and normalization prediction, a closure test is performed in simulation. Figure 7.47 demonstrates closure tests for muons and electrons for both tracker phases. The stack represents standard model background for opposite-charge analysis selection lepton-track pair (oc), while the orange line represents same-charge lepton-track pair (sc). Both are weighted to represent each phase's luminosity. In the ratio panel, which shows the ratio between opposite-charge to same-charge backgrounds for each bin, a shape agreement is demonstrated which supports the assumptions made above.

After establishing that the method can be used to correctly predict the background, a data-driven normalization factor, which is computed as the ratio between opposite-charge to same-charge data event count in the CR of $\text{BDT} < 0$. The final prediction in the SRs then becomes the same-charge data event count in the SR multiplied by the normalization factor. The computed normalization factor for phase 0 (2016) is 1.12 ± 0.044 (1.037 ± 0.05) for muons (electrons), and for phase 1 (2017-2018) is 1.066 ± 0.024 (1.049 ± 0.03) for muons (electrons). The relative errors on the normalization factors are between 2% to 5%.

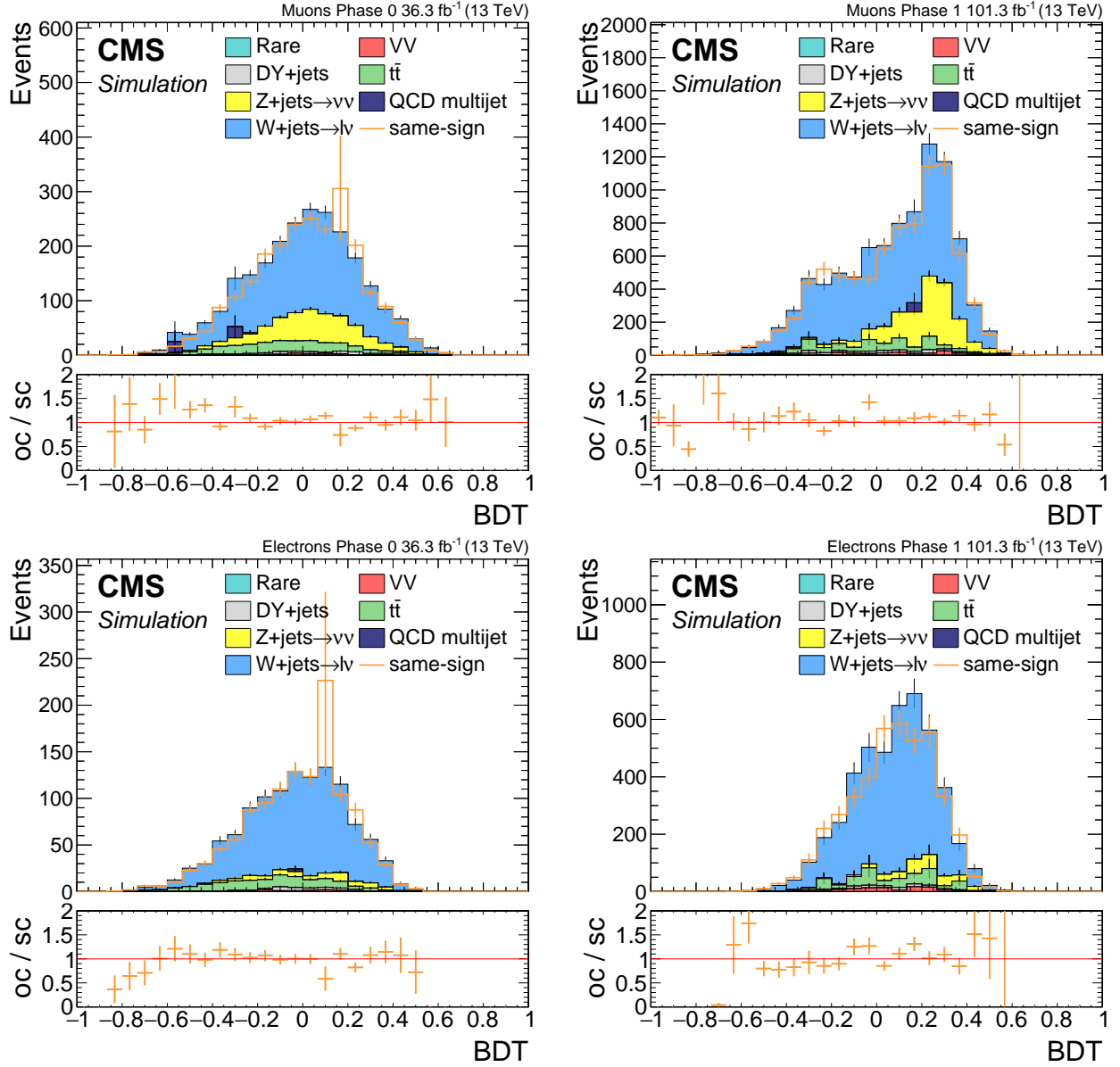


Figure 7.47: Exclusive track category closure tests for muons (top) and electrons (bottom) for phase 0 (left) and phase 1 (right). The stack represents standard model background for opposite-charge analysis selection lepton-track pair (oc), while the orange line represents same-charge lepton-track pair (sc). Both are weighted to represent each phase's luminosity. The ratio panel shows the ratio between opposite-charge to same-charge backgrounds for each bin.

7.11 Data control region plots

7.12 Signal regions definition and optimization

The signal regions in this analysis are bins in the BDT output distributions having BDT output values greater than zero. This type of search is sometimes referred to as a *shape analysis*. The significance is computed in each bin of the BDT output, and then the different values are combined statistically to produce a single significance value. One must, therefore, decide on the binning of the BDT output distribution. Since the most important merit in a beyond the standard model search such as this one is the significance, it is desired to find a binning that maximizes the significance, or at least close to the supremum.

Because of the nature of how BDTs are trained, as the BDT value increases, so does the signal counts, whether the background counts decrease. This means that the most significant bin will be closer to the right end of the distribution. The problem of finding an ideal binning is not so straight forward as one might think, since the distributions are not smooth, but are made of statistical events with potentially very low statistics. The first step in defining the SRs is defining the right most deviation that becomes the left edge of the most sensitive bin stretching all way up to the maximum BDT output value of 1. The way it is performed is by first defining a step size ε . Then, a series of computations is performed, whereby in each step i , a significance is computed for a bin of size $i \cdot \varepsilon$, i.e., in the interval $[1 - i \cdot \varepsilon, 1]$. One can then pick the left bin by taking the maximum of the series of values resulting in the previous step.

There are few points that need to be addressed. The first is what measure to choose for estimating the significance. Since the final significance, combination, and exclusion limit is calculated using via the Higgs combination tool [18], here an estimate for the purposes of optimization is used which is reviewed in [19, 20], which is referred to as the Z-value. The Z-value is a communication of the p -value by means of specifying the corresponding number of standard deviations in a one-tailed test of a Gaussian (normal) variate:

$$Z = \Phi^{-1}(1 - p) = -\Phi^{-1}(p). \quad (7.17)$$

Given the number of signal events count \hat{s} , background events count \hat{b} and its corresponding error $\delta\hat{b}$, an estimator for the significance is given by

$$Z = \frac{\hat{s}}{\sqrt{\hat{b} + \delta\hat{b}^2}}. \quad (7.18)$$

The background event count is estimated using the data-driven methods described in 7.10.2. They all involve counting events in a sideband and multiplying them by a transfer factor computed in a control region:

$$\hat{b} = N_{\text{sideband}}^{\text{SR}} \cdot \text{TF} \quad (7.19)$$

where the transfer factor TF is given by

$$\text{TF} = \frac{N_{\text{main band}}^{\text{CR}}}{N_{\text{sideband}}^{\text{CR}}} \quad (7.20)$$

The error propagation formula yields

$$\left(\frac{\delta\hat{b}}{\hat{b}} \right)^2 = \left(\frac{\delta N_{\text{sideband}}^{\text{SR}}}{N_{\text{sideband}}^{\text{SR}}} \right)^2 + \left(\frac{\delta \text{TF}}{\text{TF}} \right)^2, \quad (7.21)$$

which results in

$$\delta\hat{b}^2 = \hat{b}^2 \left[\left(\frac{\delta N_{\text{sideband}}^{\text{SR}}}{N_{\text{sideband}}^{\text{SR}}} \right)^2 + \left(\frac{\delta \text{TF}}{\text{TF}} \right)^2 \right]. \quad (7.22)$$

The second point that needs to be addressed is which signal point or points are chosen to be optimised. Since each model point has yields a different signal event count \hat{s} , they will also produce different significance values. The final values chosen maximize a range of signal points on the edge of the exclusion limit.

The third and last point to address is what step size ε to choose. If a step size too small is picked, as the BDT values are being scanned downwards from 1, there will be steps where no events are encountered either in the signal or the background due to the low statistics nature of the problem. Therefore, each signal event added will create a step upwards in the significance, while each background event will reduce it. This procedure will produce a very fine-tuned value which is not believed to be maximizing the significance, but rather it will produce wildly different results given a different set of events (from a different simulation set for example). To avoid overtraining, a step size of $\varepsilon = 0.5$ was chosen since its behavior is regular in the sense that every step adds both signal and background events, while also maintaining granularity.

Lastly, after the most significant bin has been fixed, the remaining BDT range from 0 to the left edge of the significant bin is divided equally in order to pick up any sensitivity that might remain in those bins. For the dimuon category, the bin width is chosen as 0.1, while for the exclusive track categories, it is 0.05. The final signal regions are as follows:

Table 7.11: Signal Regions

Category	Flavor	Phase	SR	Signal Regions
Dilepton	Muons	all	6	$[0, 0.1, 0.2, 0.3, 0.4, 0.5, 1]$
Exclusive Track	Muons	0	13	$[0, 0.05, 0.1, 0.15, 0.2, \dots, 0.5, 0.55, 0.6, 1]$
Exclusive Track	Muons	1	12	$[0, 0.05, 0.1, 0.15, 0.2, \dots, 0.5, 0.55, 1]$
Exclusive Track	Electrons	all	11	$[0, 0.05, 0.1, 0.15, 0.2, \dots, 0.5, 1]$

7.13 Systematic uncertainties

7.14 Data quality

7.15 Results and interpretation

Chapter 8

Summary

Chapter 9

Latex stuff

9.1 Some examples

9.1.1 Multiline comment

This is a line in introduction.

9.1.2 Fixme note

This is the introduction to the thesis. **FiXme Note: This is a fixme note** **FiXme Note: what** **FiXme Note: WHAT THE HELL AFTER**

9.1.3 Tables

h - here t - top b - bottom p - special page ! - even if not pretty

Table 9.1: Table captions are above the table whereas figure captions are below.

Parameter	Value 1	Value 2
s	10.0	20.0
t	20.0	30.0
u	30.0	40.0

9.1.4 Cross References

9.1.4 81 section 9.1.4

9.1.5 Particles

Hello World $\tilde{\chi}_1^0 \pi \eta_c$ GeV E_T^{miss} hey GeV E_T^{miss} π new one $\tilde{\chi}_1^0 \tilde{\chi}_1^0$

9.1.6 Citing

[1] SOS analysis

9.1.7 Glossary

Using glossary for computer computer plural form computers upper case first Computer upper case first plural Computers. To use for symbol π

9.1.8 Acronyms

First use of acronym SOS and second SOS. You can reset this and do again Soft-Opposite-Sign (SOS) and second time again SOS. Long version Soft-Opposite-Sign. Full version Soft-Opposite-Sign (SOS). Short version SOS.

9.1.9 fractions

Here's a fraction with a slash:

$$\text{Track Isolation}_\ell = \sum_{\substack{\text{tracks from PV} \\ \text{in } \Delta R < 0.3}} p_{\text{T}} / p_{\text{T}}(\ell) \quad (9.1)$$

List of Corrections

Note: fill in signal model stuff	13
Note: make sure we define the different deltaM somewhere	13
Note: Add description here.	59
Note: Add jetty/tautau simulation plot	60
Note: This is a fixme note	81
Note: what	81
Note: WHAT THE HELL	81

Glossary

computer is a programmable machine that receives input, stores and manipulates data, and provides output in a useful format. 73

$\Delta\mathbf{m}$ mass difference between electroweakinos. If not explicitly said otherwise, it is the mass difference between $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$, i.e., $\Delta\mathbf{m} = \Delta\mathbf{m}^0 = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$. 27, 31

$\Delta\mathbf{m}^0$ mass difference between the neutral electroweakinos $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$, i.e., $\Delta\mathbf{m}^0 = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$. 26

ΔR separation. 19, 23, 24

η pseudorapidity. 23

$E_{\mathbf{T}}^{\text{miss}}$ add description. 15, 16

$H_{\mathbf{T}}^{\text{miss}}$ add description. 15, 16, 50

$m_{\ell\ell}$ invariant mass. 19–21

neutralino add description. 15, 16

ϕ azimuthal angle measured in radians. 23

π ratio of circumference of circle to its diameter. 73

$p_{\mathbf{T}}$ transverse momentum. 15, 18, 19, 23, 27, 31

Acronyms

BDT Boosted Decision Tree. 31, 41, 43, 51–53, 55, 58–60, 63

CERN European Organization for Nuclear Research. 9

CMS Compact Muon Solenoid. 9, 45, 65

CR Control Region. 62–65

CSV Combined Secondary Vertex. 40

DM Dark Matter. 15

ECAL Electromagnetic Calorimeter. 27

ISR Initial State Radiation. 15, 16, 26, 27, 50, 51

LHC Large Hadron Collider. 9

LSP Lightest SUSY Particle. 15

MC Monte Carlo. 35, 36, 46, 52, 59, 62, 64, 65

PF Particle Flow. 40

POG Physics Object Group. 35, 36

PU Pile-Up. 40, 45

QCD Quantum Chromodynamics. 15, 16, 50, 59

SF Scale Factor. 35

SM Standard Model. 15, 16, 23, 24, 33, 50–52, 58, 62

SOS Soft-Opposite-Sign. 21–24, 51, 74

SR Signal Region. 52, 58–60, 62–64

SUSY Supersymmetry. 15

WIMP Weakly Interacting Massive Particle. 15

Bibliography

- [1] CMS Collaboration, A. Tumasyan *et al.*, “Search for supersymmetry in final states with two or three soft leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **04** (2022) 091, arXiv:2111.06296 [hep-ex].
- [2] https://gitlab.cern.ch/cms-muonP0G/MuonReferenceEfficiencies/-/tree/master/EfficienciesStudies/2016_legacy_rereco/Jpsi.
- [3] <https://indico.cern.ch/event/791428/contributions/3287172/attachments/1781977/2899551/2016LegacyReReco2.pdf>.
- [4] BRIL Group, “Bril work suite.”
<http://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html>, 2017.
- [5] https://lguzzi.web.cern.ch/lguzzi/Tau3Mu/muonP0G_mediumID_dR.pdf.
- [6] T. C. collaboration, “Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ tev,” *Journal of Instrumentation* **10** no. 02, (Feb, 2015) P02006–P02006. <https://doi.org/10.1088%2F1748-0221%2F10%2F02%2Fp02006>.
- [7] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm,” *Journal of High Energy Physics* **2008** no. 04, (Apr, 2008) 063–063.
<https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063>.
- [8] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay, M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla, “Tmva - toolkit for multivariate data analysis,” 2007.
<https://arxiv.org/abs/physics/0703039>.
- [9] <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>.
- [10] S. Chatrchyan, , V. Khachatryan, and A. M. et. al, “Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC,” *Journal of High Energy Physics* **2011** no. 6, (Jun, 2011) .
<https://doi.org/10.1007%2Fjhep06%282011%29077>.
- [11] K. Rehermann and B. Tweedie, “Efficient identification of boosted semileptonic top quarks at the LHC,” *Journal of High Energy Physics* **2011** no. 3, (Mar, 2011) .
<https://doi.org/10.1007%2Fjhep03%282011%29059>.
- [12] R. Ellis, I. Hinchliffe, M. Soldate, and J. Van Der Bij, “Higgs decay to $\tau^+\tau^-$ -a possible signature of intermediate mass higgs bosons at high energy hadron colliders,” *Nuclear Physics B* **297** no. 2, (1988) 221–243.
<https://www.sciencedirect.com/science/article/pii/0550321388900193>.

- [13] **ATLAS** Collaboration, G. Aad *et al.*, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics,” [arXiv:0901.0512 \[hep-ex\]](#).
- [14] **CMS** Collaboration, G. L. Bayatian *et al.*, “CMS technical design report, volume II: Physics performance,” *J. Phys. G* **34** no. 6, (2007) 995–1579.
- [15] Z. Han, G. D. Kribs, A. Martin, and A. Menon, “Hunting quasidegenerate higgsinos,” *Physical Review D* **89** no. 7, (Apr, 2014) .
<https://doi.org/10.1103/PhysRevD.89.075007>.
- [16] H. Baer, A. Mustafayev, and X. Tata, “Monojet plus soft dilepton signal from light higgsino pair production at LHC14,” *Physical Review D* **90** no. 11, (Dec, 2014) .
<https://doi.org/10.1103/PhysRevD.90.115007>.
- [17] A. Barr and J. Scoville, “A boost for the EW SUSY hunt: monojet-like search for compressed sleptons at LHC14 with 100 fb⁻¹,” *Journal of High Energy Physics* **2015** no. 4, (Apr, 2015) . [https://doi.org/10.1007/JHEP04\(2015\)29147](https://doi.org/10.1007/JHEP04(2015)29147).
- [18] <https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/>.
- [19] L. Demortier, “P values and nuisance parameters,” in *Statistical issues for LHC physics. Proceedings, Workshop, PHYSTAT-LHC, Geneva, Switzerland, June 27-29, 2007*, p. 23. 2008. <http://cds.cern.ch/record/1099967/files/p23.pdf>.
- [20] R. D. Cousins, J. T. Linnemann, and J. Tucker, “Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process,” *Nucl. Instrum. Meth. A* **595** (2008) 480, [arXiv:physics/0702156 \[physics.data-an\]](#).

List of Figures

7.1	Signal Models Feynman Diagrams	13
7.2	Signal E_T^{miss} and H_T^{miss} distributions	16
7.3	Signal <i>number of jets</i> and <i>leading jet</i> p_T distributions	17
7.4	Signal <i>number of b-tagged jets</i> distributions	17
7.5	Signal $\min \Delta\phi(E_T^{\text{miss}}, \text{jets})$ and $\min \Delta\phi(H_T^{\text{miss}}, \text{jets})$ distributions	18
7.6	Signal p_T distributions	19
7.7	Signal p_T distributions split into barrel and endcaps	20
7.8	Signal $ \eta $ distributions	20
7.9	Signal generator level $m_{\ell\ell}$ distributions	21
7.10	Signal $m_{\mu\mu}$ vs. p_T	22
7.11	Signal reconstructed $m_{\mu\mu}$	22
7.12	Signal generator level ΔR distributions	23
7.13	Signal $\Delta R(\mu\mu)$ vs. p_T	24
7.14	Signal reconstructed $\Delta R(\mu\mu)$	24
7.15	Spatial separation between reconstructed electrons and the leading jet $\Delta R(j_1, e)$	26
7.16	p_T distribution of reconstructed electrons with loose ID	27
7.17	$ \eta $ distribution of reconstructed electrons with loose ID	27
7.18	medium and tight ID working points distribution of reconstructed electrons	28
7.19	standard isolation and jet-isolation distribution of reconstructed electrons	29
7.20	$ \eta $ distribution of reconstructed electrons with loose ID passing jet-isolation	29
7.21	Spatial separation between reconstructed muons and the leading jet $\Delta R(j_1, \mu)$	31
7.22	Reconstructed muons p_T	32
7.23	$ \eta $ distribution of reconstructed muons with loose ID before and after $p_T > 2 \text{ GeV}$ cut	32
7.24	medium ID working point distribution of reconstructed muons	33
7.25	tight ID working point distribution of reconstructed muons	34
7.26	jet-isolation distribution of reconstructed muons	34
7.27	Simulation barrel muons fits	37
7.28	Simulation endcaps muons fits	37
7.29	Data barrel muons fits	38
7.30	Data endcaps muons fits	38
7.31	Efficiencies and scale factors	39
7.32	Muon track BDT training inputs	42
7.33	Track BDT output plots	43
7.34	Track BDT ROC curves	44
7.35	Dimuon BDT output and ROC curve	53
7.36	dimuon training input distribution	54
7.37	Exclusive track category BDT outputs	55
7.38	Exclusive track category ROC curve	56
7.39	exclusive track training input distribution	57

7.40	Dimuon simulation BDT output	60
7.41	Dimuon simulation BDT output jetty tautau	60
7.42	Dimuon simulation BDT inputs	61
7.43	Exclusive track simulation BDT output	63
7.44	Exclusive track plus muon simulation BDT inputs	64
7.45	Dimuon 2017 jetty background closure plot	67
7.46	Ditau invariant mass distributions	70
7.47	Exclusive track category closure tests	72

List of Tables

7.1	Base selection applied to all analysis categories	18
7.2	Generator level efficiency on muons selections	21
7.3	Selection criteria for Tags and Probes	36
7.4	Track BDT input variables	42
7.5	Signal Efficiency	46
7.6	Background Efficiency	47
7.7	Transfer Factor	47
7.8	Significance $s/\sqrt{b + \epsilon_b^2}$	48
7.9	Dimuon BDT input variables	53
7.10	Exclusive track BDT input variables	56
7.11	Signal Regions	75
9.1	Table captions are above the table whereas figure captions are below.	81