

Search for compressed mass Higgsino production with low-momentum lepton tracks with the CMS experiment

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



Abstract This is the abstract

Contents

1	Introduction	1
2	Theoretical background	3
2.1	Principle of Least Action	3
2.2	The Quantum	4
2.2.1	Annihilation and Creation Operators	5
2.3	Relativistic Quantum Field Theory	6
2.3.1	Attempts at Relativistic Quantum Mechanics	7
2.3.2	Classical Field Theory	8
2.3.3	Quantization	9
2.3.3.1	Canonical Quantization	9
2.3.3.2	Path Integrals	10
2.3.4	Interactions	10
2.3.4.1	The Cross Section and Decay Rate	10
2.3.4.2	Interacting Fields	11
2.3.5	Feynman Diagrams	11
2.3.6	Renormalization	11
2.4	Symmetries	13
2.4.1	Conservation Laws	14
2.4.2	Noether's Theorem	15
2.4.3	Groups	16
2.4.4	Symmetries of the Standard Model	16
2.4.4.1	Parity	16
2.5	The Standard Model	16
2.5.1	Shortcomings of the Standard Model	16
2.6	Supersymmetry	16
2.7	Motivational Discussion for Supersymmetry	16
2.7.1	No go theorem	16
2.7.2	Dark Matter Candidate	16
2.7.3	Naturalness	16
3	Experimental setup	17
3.1	The Large Hadron Collider	17
3.2	The Compact Muon Solenoid experiment	18
4	Search for compressed Higgsinos with soft lepton tracks	21
4.1	Signal models	21
4.2	Previous searches	21
4.3	Search strategy	22
4.3.1	Final state with two muons	22
4.3.2	Final state with one lepton and one track	22

4.4	Signal signature and base selection	23
4.4.1	Missing transverse energy	23
4.4.2	Jets and hardronic activity	23
4.4.3	Base selection	24
4.4.4	Lepton kinematics	27
4.4.4.1	Lepton η and transverse momentum p_T	27
4.4.4.2	Invariant mass $m_{\ell\ell}$	30
4.4.4.3	Lepton separation ΔR	31
4.4.5	Main drivers of sensitivity	34
4.5	Simulated samples	35
4.5.1	Standard Model simulated samples	35
4.5.2	Signal simulated samples	35
4.6	Object definition and selection	36
4.6.1	Electrons	36
4.6.2	Muons	40
4.6.3	Scale factors	44
4.6.4	Missing transverse energy	49
4.6.5	Jets	49
4.6.6	Tracks and multivariate selection	50
4.6.7	Isolation	54
4.7	Trigger	58
4.8	Event selection	59
4.8.1	Preselection	59
4.8.2	Category selection	60
4.8.2.1	Dilepton selection	60
4.8.2.2	Exclusive track selection	60
4.8.3	Binary event classifier	61
4.8.3.1	Dimuon category	61
4.8.3.2	Exclusive track category	64
4.9	Characterization and estimation of the Standard Model backgrounds	67
4.9.1	Characterization of the Standard Model backgrounds	67
4.9.1.1	Dimuon category	69
4.9.1.2	Exclusive track category	71
4.9.2	Estimation of the Standard Model backgrounds	73
4.9.2.1	Jetty background estimation	74
4.9.2.2	Ditau Drell-Yann background estimation	76
4.9.2.3	Exclusive track background estimation	78
4.10	Data control region plots	80
4.11	Optimization of search bins	83
4.12	Systematic uncertainties	85
4.12.1	Data driven transfer factors	85
4.12.2	Uncertainty in jetty background template	85
4.12.3	Uncertainty in $Z/\gamma^* \rightarrow \tau^-\tau^+$ background	86
4.13	Data quality aspects and background validation	88
4.13.1	L1 prefire issue in 2016 and 2017 data	88
4.13.2	EE noise in 2017 data	88
4.13.3	HEM failure in 2018 data	88
4.13.4	Validation in same-sign CR	88
4.14	Results and interpretation	90

CONTENTS

5	Summary	91
6	Latex stuff	93
6.1	Some examples	93
6.1.1	Multiline comment	93
6.1.2	Fixme note	93
6.1.3	Tables	93
6.1.4	Cross References	93
6.1.5	Particles	93
6.1.6	Citing	93
6.1.7	Glossary	93
6.1.8	Acronyms	94
6.1.9	fractions	94
6.1.10	Quarantine	94
	Glossary	97
	Acronyms	99

CONTENTS

Chapter 1

Introduction

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

Chapter 2

Theoretical background

The search presented in this thesis is for a simplified model of a supersymmetric theory. The amount of theoretical background required to fully understand this theory can fill many books, and therefore, this document does not attempt to give an exhaustive description of it. Instead, I attempt to give a brief tour of topics that contribute to the understanding the subject, which are also of personal interest. Especially I wanted to also explore the theoretical motivation for supersymmetry, alongside a contribution to the philosophical discussion that normally accompanies such arguments. Whenever possible, I pick a description of a concept that I find intriguing and inspiring in a way that reminds me of my initial spark and inspiration for perusing a PhD in physics. Good sources for these topics are [1, 2].

2.1 Principle of Least Action

The earliest formulation of classical mechanics is normally attributed to the works of Sir Isaac Newton from the 17th century, which is also referred to as Newtonian mechanics. It is based on the then-newly developed mathematics of calculus. A central theorem in calculus is Fermat's theorem, which states that if a function has a local extremum at some point and is differentiable there, then the function's derivative at that point must be zero. The equation of motion is given by Newton's second law, which is an ordinary differential equation given by:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d(m\mathbf{v})}{dt}. \quad (2.1)$$

When the mass m is constant, this is equivalent to the famous $\mathbf{F} = m\mathbf{a}$. In modern physics, a more generalized approach is used based on an *action*. It has been developed in the 18th centenary, and is able to reproduce Newtonian mechanics, but also able to generalize to handle Quantum Mechanics (QM), Relativistic Quantum Field Theory (RQFT) and even General Relativity (GR). The development of that principle was carried in different times by different people, and can be formulated in equivalent manners. In RQFT it is useful to use a *Lagrangian* and therefore it will be shown here, rather than the *Hamiltonian* formulation. The two formulations are equivalent, however. Given N generalized coordinates $\mathbf{q} = (q_1, q_2, \dots, q_N)$, a *Lagrangian* of the system is written $L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)$. In non-relativistic mechanics for a system of particles in the absence of a magnetic field $L = T - V$ where T is the total kinetic energy of the system and V is the potential energy of the system. For other systems, writing a Lagrangian is not straight forward, and we'll just assume that it is given. The *action* of the system is a functional of the N generalized coordinates, denoted \mathcal{S} , given by:

$$\mathcal{S}[\mathbf{q}, t_1, t_2] = \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) dt \quad (2.2)$$

where the dot denotes the time derivative, and t is time. The principle of least action is then:

The path taken by the system between times t_1 and t_2 and configurations q_1 and q_2 is the one for which the action is stationary (no change) to first order.

Mathematically, that is equivalent to requiring $\delta\mathcal{S} = 0$ or:

$$\delta \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) dt = 0. \quad (2.3)$$

The principle of least action has been preceded by earlier ideas in optics, such as that for the path of light reflecting from a mirror, the angle of incidence equals the angle of reflection. The principle of least action is the variational equivalent in the calculus of variations of Fermat's theorem in calculus. It is used in order to find a path that extremizes the Lagrangian. Interestingly enough, Fermat also coined Fermat's principle, which states that "light travels between two given points along the path of shortest time", which is an earlier example of the principle of least action. Using this principle, one can arrive at the equations of motion of the system. For a classical system, those would be equivalent to Newton's laws of motion Eq. 2.1. Solving Eq. 2.3, one arrives at Euler–Lagrange equations:

$$\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} = 0. \quad (2.4)$$

Solving Euler–Lagrange equations gives the equations of motion of the system. In field theory, an analogous equation is used to calculate the dynamics of a field.

2.2 The Quantum

The main object that is the subject of research in particle physics is, of course, a particle. More precisely, an elementary particle or fundamental particle, which is a subatomic particle that is not composed of other particles. The electron is an example of such a fundamental particle, which was also the first to be discovered by Thomson in 1897. The description and properties of the particles have radically evolved over time, and so did the mathematical language that is used to describe them. In classical electromagnetism for example, one can use abstractions such as a point charge, point mass or electron point by using a Dirac delta function δ in the charge and mass distributions. In quantum mechanics, a wave function $\Psi(\mathbf{x}, t)$, that assigns a complex number to each point \mathbf{x} at each time t , is used. The wave function is governed by the Schrödinger equation [3–5]. The time-dependent Schrödinger equation is:

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle. \quad (2.5)$$

For a single nonrelativistic particle in one dimension that becomes:

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, t) \right] \Psi(x, t). \quad (2.6)$$

The parameter m is the mass of the particle, and $V(x, t)$ is the potential that represents the environment in which the particle exists. This can be easily generalized to include more than one particle. However, nonrelativistic quantum mechanics has a shortcoming, in that the Schrödinger equation for massive particles has a fixed number of particles governing the state of the system. It is not surprising given the fact that in classical mechanics, and therefore nonrelativistic quantum mechanics by extension, mass is never created nor destroyed. In order

to accommodate the observation that particles are being created and destroyed, a relativistic treatment is needed. That is the goal of RQFT.

But the equivalent of particles DO actually arise in nonrelativistic quantum mechanics: when they are massless. In fact, the formalism for creating and destroying massless particles, quanta, is generalized from quantum mechanics to RQFT. The quantum arises in the quantum mechanical harmonic oscillator. Classically, a harmonic oscillator is a system that, when displaced from its equilibrium position, experiences a restoring force F proportional to the displacement x :

$$\mathbf{F} = -k\mathbf{x}, \quad (2.7)$$

where k is a positive constant. The potential energy stored in a simple harmonic oscillator at position x is:

$$U = \frac{1}{2}kx^2. \quad (2.8)$$

Writing a Hamiltonian and promoting the observables to operators we get:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2, \quad (2.9)$$

where m is the particle's mass, k is the force constant, $\omega = \sqrt{k/m}$ is the angular frequency of the oscillator, \hat{x} is the position operator, and \hat{p} is the momentum operator. Solving the time-independent Schrödinger equation gives the energy levels

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) = (2n + 1)\frac{\hbar}{2}\omega. \quad (2.10)$$

It is interesting to note that the energies are quantized and equally spaced with discrete energy values of integer-plus-half multiples of $\hbar\omega$.

2.2.1 Annihilation and Creation Operators

We define ladder operators

$$\begin{aligned} \hat{a} &\equiv \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i\hat{p}}{m\omega_0} \right) \\ \hat{a}^\dagger &\equiv \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{i\hat{p}}{m\omega_0} \right). \end{aligned} \quad (2.11)$$

As can be seen, \hat{a} is not Hermitian. Using $[\hat{x}, \hat{p}] = i\hbar$ it is easy to show that

$$\begin{aligned} [\hat{a}, \hat{a}^\dagger] &= 1 \\ \hat{a}\hat{a}^\dagger &= 1 + \hat{a}^\dagger\hat{a}. \end{aligned} \quad (2.12)$$

By reversing 2.11 we get

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}} (\hat{a} + \hat{a}^\dagger), \quad \hat{p} = i\sqrt{\frac{\hbar m\omega}{2}} (\hat{a} - \hat{a}^\dagger) \quad (2.13)$$

and the Hemiltonian becomes

$$\hat{H} = \hbar\omega_0 \left(\hat{a}^\dagger\hat{a} + \frac{1}{2} \right) \equiv \hbar\omega_0 \left(\hat{N} + \frac{1}{2} \right). \quad (2.14)$$

Finding eigenvalues for \hat{H} becomes finding eigenvalues of the *number operator* $\hat{N} \equiv \hat{a}^\dagger \hat{a}$, which are

$$\hat{N} |n\rangle = n |n\rangle. \quad (2.15)$$

Operating with the ladder operators on the energy eigenstates gives

$$\begin{aligned}\hat{a}^\dagger |n\rangle &= \sqrt{n+1} |n+1\rangle \\ \hat{a} |n\rangle &= \sqrt{n} |n-1\rangle.\end{aligned}\quad (2.16)$$

It is seen that \hat{a}^\dagger , in essence, appends a single quantum of energy to the oscillator, while \hat{a} removes a quantum. Furthermore, acting with the number operator \hat{N} yields

$$\begin{aligned}\hat{N} \hat{a}^\dagger |n\rangle &= (n+1) \hat{a}^\dagger |n\rangle \\ \hat{N} \hat{a} |n\rangle &= (n-1) \hat{a} |n\rangle.\end{aligned}\quad (2.17)$$

Due to this, \hat{a} is called a annihilation operator ("lowering operator"), and \hat{a}^\dagger a creation operator ("raising operator"). The two operators together are called ladder operators. In quantum field theory, these operators destroy and create particles, which correspond here to a quanta of energy of $\hbar\omega$.

2.3 Relativistic Quantum Field Theory

In the first quarter of the twentieth century, two of the most successful theories in modern physics have been developed, special relativity and quantum mechanics. Special relativity was necessary to solve the incompatibility of Maxwell's equations of electromagnetism with Newtonian mechanics. In addition, experimentally, the Michelson–Morley experiment's null result demonstrated that the historically hypothesized luminiferous aether did not exist. Special relativity diverges from classical mechanics in high-velocities. Quantum mechanics, on the other hand, arose gradually from theories to explain observations that could not be reconciled with classical physics, such as Max Planck's solution to the black-body radiation problem, and the correspondence between energy and frequency in Albert Einstein's photoelectric effect. Quantum mechanics differs from classical physics in that energy, momentum, angular momentum, and other quantities of a bound system are restricted to discrete values; objects have characteristics of both particles and waves; and there are limits to how accurately the value of a physical quantity can be predicted prior to its measurement, given a complete set of initial conditions (the uncertainty principle).

Since classical mechanics diverged into two different directions, namely, quantum mechanics and special relativity (which later on developed further into general relativity, but that's beyond the concern here), it was clear that a theory that incorporates both developments is needed. The first effort came from an attempt in creating a quantum mechanical theory of the electromagnetic field. It was also crucial to develop a theory, in which the number of particles changes, in order describe processes such as a β -decay or the emission of a photon by an electron dropping into a quantum state of lower energy in an atom.

Quantum field theory successfully combines classical field theory, special relativity, and quantum mechanics. QFT treats particles as excited states (also called quanta) of their underlying quantum fields, which are more fundamental than the particles. The equation of motion of the particle is determined by minimization of the Lagrangian, a functional of fields associated with the particle. Interactions between particles are described by interaction terms in the Lagrangian involving their corresponding quantum fields. Each interaction can be visually represented by Feynman diagrams according to perturbation theory in quantum mechanics.

2.3.1 Attempts at Relativistic Quantum Mechanics

At first glance, fields are not the only way to try and reconcile quantum mechanics and relativity. A naive attempt [2] could be to take the Schrödinger equation 2.5 and write a Hamiltonian in a relativistic notion $H = \sqrt{\hat{\mathbf{p}}^2 + m^2}$ (taking as usual $\hbar = c = 1$). Plugging it as is into the Schrödinger equation will result in the time derivative outside the square root, while the space derivatives under it, which is not in the spirit of relativity. Squaring the differential operators before applying them to the wave function and collecting terms results in the *Klein-Gordon equation*:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2 \right) \Psi(\mathbf{x}, t) = 0. \quad (2.18)$$

It is second-order in both space and time derivatives, and they appear in a symmetric fashion. The Ψ in the equation is the usual quantum mechanical wave function. There are two problems with sticking to the wave function. The first, is that the norm of a state $\langle \Psi, t | \Psi, t \rangle$ is not in general time independent. Thus probability is not conserved. The Klein-Gordon equation obeys relativity, but not quantum mechanics. This specific problem is solved (for spin-one-half particles) by the *Dirac equation*. In its original form written by Dirac [6]:

$$\left(\beta mc^2 + c \sum_{n=1}^3 \alpha_n p_n \right) \Psi(\mathbf{x}, t) = i\hbar \frac{\partial \Psi(\mathbf{x}, t)}{\partial t} \quad (2.19)$$

where $\Psi(\mathbf{x}, t)$ again is to be interpreted as an ordinary quantum mechanical wave function for the electron of rest mass m with spacetime coordinates x, t . The p_1, p_2, p_3 are the components of the momentum, understood to be the momentum operator in the Schrödinger equation. The new elements in this equation are the four 4×4 matrices $\alpha_1, \alpha_2, \alpha_3$ and β , and the four-component wave function Ψ . There are four components in Ψ because the evaluation of it at any given point in configuration space is a bispinor. It is interpreted as a superposition of a spin-up electron, a spin-down electron, a spin-up positron, and a spin-down positron. The 4×4 matrices α_k and β are all Hermitian and satisfy:

$$\alpha_i^2 = \beta^2 = I_4, \quad (2.20)$$

and they all mutually anticommute:

$$\begin{aligned} \alpha_i \alpha_j + \alpha_j \alpha_i &= 0 \quad (i \neq j) \\ \alpha_i \beta + \beta \alpha_i &= 0. \end{aligned} \quad (2.21)$$

It turns out that the Dirac equation is fully consistent with relativity provided. There are, however, some problems. The minimum size of the matrices of 4×4 imply extra two "spin" states. They also imply negative eigenvalues for the Hamiltonian which indicate that there is no ground state. Dirac postulated his famous *Dirac sea* of electrons to suggest that the negative energy states are all occupied. An electron in the sea could then be excited to a positive energy state leaving behind a *hole* in the Dirac sea. This hole would appear to have positive charge, and positive energy. Dirac therefore predicted (in 1927) the existence of the positron, a particle with the same mass as the electron, but opposite charge. The positron was found experimentally five years later.

The problem with this solution though, is that we've started by trying to describe a theory of a single half-spin particle, and ended up describing a theory with infinite amount of particles. Even if this is taken to be satisfactory, this theory still cannot describe particles that do not obey Pauli exclusion, such as photons or pions. The problem lies in the difference between the way that nonrelativistic quantum mechanics and special relativity treats space and time. In special relativity, space and time are treated on equal footing. In nonrelativistic quantum

mechanics, however, space is an operator, while time isn't. It turns out that turning time into an operator is a very difficult problem. The approach that proved to be fruitful is to make space a *label*, just as time is, by turning the wave function Ψ into a *field*. Space and time are now labels in a *quantum field* $\varphi(\mathbf{x}, t)$ of operators. Each point in space and time now point to an operator. This allows one to really treat space and time on an equal footing.

2.3.2 Classical Field Theory

After the naive attempts at a relativistic quantum mechanics introduced in Section 2.3.1, two successful and widely used methods of constructing quantum field theories are described in Section 2.3.3. The first is the canonical quantization in Section 2.3.3.1, and the second is the path integrals formalism in Section 2.3.3.2. They involve starting from a classical field theory and quantizing it. In a classical field theory, the equation of motion can be derived from variation of an action $\mathcal{S} = \int dt L$, where L is the Lagrangian, which is the spatial integral of a Lagrangian density \mathcal{L} , so that $L = \int d^3x \mathcal{L}$. The Lagrangian density is a function of one or more fields $\phi(x)$, and their derivatives $\partial_\mu \phi$, so that

$$\mathcal{S} = \int dt L = \int \mathcal{L}(\phi, \partial_\mu \phi) d^4x. \quad (2.22)$$

Following the principle of least action described in Section 2.1, we take the action \mathcal{S} to an extremum and write

$$\begin{aligned} 0 &= \delta \mathcal{S} \\ &= \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta \phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta (\partial_\mu \phi) \right\} \\ &= \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta \phi - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \delta \phi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta \phi \right) \right\}. \end{aligned} \quad (2.23)$$

The last term can be turned into a surface integral over the boundary of the region of integration. Since $\delta \phi$ vanish on the spatial boundary, then the surface term is zero. After rearrangement, we arrive at the Euler-Lagrange equation of motion for a field,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (2.24)$$

If the Lagrangian contains more than one field, there is one such equation for each. The Hamiltonian of a discrete system can be written as

$$H \equiv \sum p \dot{q} - L, \quad (2.25)$$

where q is a dynamical variable, and $p \equiv \partial L / \partial \dot{q}$ is the conjugate momentum. To generalize to continuous system we define the *momentum density* conjugate to $\phi(\mathbf{x})$ as

$$\pi(\mathbf{x}) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{x})}, \quad (2.26)$$

and the Hamiltonian can be expressed, using the Hamiltonian desnsity \mathcal{H} as:

$$H = \int d^3x [\pi(\mathbf{x}) \dot{\phi}(\mathbf{x}) - \mathcal{L}] \equiv \int d^3x \mathcal{H}. \quad (2.27)$$

As an example, consider the theory of a single real scalar field $\phi(x)$ with the Lagrangian

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2. \quad (2.28)$$

Following the usual procedure and applying the Euler-Lagrange equation gives the equation of motion

$$(\partial^\mu \partial_\mu + m^2) \phi = 0, \quad (2.29)$$

which is the well-known Klein-Gordon equation. Here, ϕ is a classical field, and not a wave function, nor a quantum field. The Hamiltonian that results from the procedure described above is

$$H = \int d^3x \left[\frac{1}{2} \pi^2 + \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 \right]. \quad (2.30)$$

2.3.3 Quantization

As described in Section 2.3.2, two methods of constructing quantum field theories are widely used. The first is the canonical quantization in Section 2.3.3.1, and the second is the path integrals formalism in Section 2.3.3.2. The path integrals formalism has an advantage in that it uses the Lagrangian formalism rather than the Hamiltonian. The Lagrangian formalism is explicitly Lorentz invariant, and in general, it is in practice easier to guess the correct form of the Lagrangian of a theory, which naturally enters the path integrals than the Hamiltonian. The advantage of the canonical quantization is that unitarity of the S-matrix is more explicit than in the path integral approach. The methods are described here in a very qualitatively manner. For an explicit mathematical formulation, Ref. [1, 2] are great sources for that.

2.3.3.1 Canonical Quantization

Canonical quantization starts with a classical field theory, and then *quantized* by promoting the dynamical variables to operators that obey canonical commutation relations. It can be demonstrated with the example of the Klein-Gordon case, which has the classical Lagrangian 2.28. Promoting the field and momentum density to operators, the commutation relations generalize to:

$$\begin{aligned} [\phi(\mathbf{x}), \pi(\mathbf{y})] &= i\delta^{(3)}(\mathbf{x} - \mathbf{y}); \\ [\phi(\mathbf{x}), \phi(\mathbf{y})] &= [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0. \end{aligned} \quad (2.31)$$

Writing the Klein-Gordon equation in Fourier space, one gets:

$$\left[\frac{\partial^2}{\partial t^2} + (|\mathbf{p}|^2 + m^2) \right] \phi(\mathbf{p}, t) = 0. \quad (2.32)$$

This is the same as the equation of motion for a simple harmonic oscillator with the frequency $\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}^2 + m^2|}$. Therefore a similar treatment to Section 2.2 can be done here. Ladders operators are introduced, only that now each Fourier mode of the field is treated as an independent oscillator with its own a and a^\dagger . The spectrum of the Klein-Gordon Hamiltonian can then be found in the same manner, and can be written as:

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \left(a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \frac{1}{2} [a_{\mathbf{p}}, a_{\mathbf{p}}^\dagger] \right). \quad (2.33)$$

The operator $a_{\mathbf{p}}^\dagger$ creates a particle with momentum \mathbf{p} and energy $\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}^2 + m^2|}$. The particles follow the proper relativistic energy-momentum relation, and have strictly positive

energy. Since $a_{\mathbf{p}}^\dagger$ and $a_{\mathbf{q}}^\dagger$ commute, two particles are interchangeable. Moreover, since arbitrarily many particles can be produced for a single mode \mathbf{p} , the particles obey *Bose-Einstein statistics*. In a theory of half-integer spin particles, anticommutators are to be used. The next steps in this formalism is to compute correlation functions, and eventually write down the full Feynman rules for the theory, in order to compute cross sections and decay rates.

2.3.3.2 Path Integrals

The path integral formalism is an alternative construction to quantum mechanics developed by Richard Feynman, and is proven to be equivalent to the wave equation of Schrödinger, and the matrix algebra of Heisenberg, Born and Jordan. It is also used as an alternative way to construct quantum field theories, as an alternative to the canonical quantization. Using this formalism, it is easier compute propagators and derive Feynman rules. It also generalizes better to non-Abelian gauge theories. Moreover, since it uses the Lagrangian, rather than the Hamiltonian, as its fundamental quantity, it explicitly preserves all symmetries of a theory. Using path integrals allows the direct computation of the scattering amplitude of a certain interaction process, rather than the establishment of operators and state spaces.

Suppose we are interested to compute the amplitude for a particle to travel from one point x_a to another x_b in a given time T . The amplitude $U(x_a, x_b, T)$ in the canonical Hamiltonian formalism using the time evolution operator is given by

$$U(x_a, x_b, T) = \langle x_b | e^{-iHT/\hbar} | x_a \rangle \quad (2.34)$$

In the path integral approach, the total time T is divided into N small intervals, and the overall amplitude is the product of the amplitude of evolution within each interval, integrated over all intermediate states. The propagation amplitude becomes

$$\langle x_b | e^{-iHT/\hbar} | x_a \rangle = U(x_a, x_b, T) = \int \mathcal{D}x(t) e^{iS[x(t)]/\hbar}, \quad (2.35)$$

where $S[x(t)]$ is the classical action, and $\int \mathcal{D}x(t)$ is another way of writing "sum over all paths". The functional formula then allows for the calculation of correlation functions, and eventually write down the full Feynman rules for the theory.

2.3.4 Interactions

The goal of every scientific theory is to make predictions about measurements. In the context of quantum field theory, it is normally one of two generic cases: one incoming particle, for which a decay rate is computed, or two incoming particles, for which a cross section is computed. For this, a recipe for computing a scattering amplitude and converting it into a measurable is needed.

2.3.4.1 The Cross Section and Decay Rate

The *cross section* is the likelihood of any particular final state from collision of two beams of particles with well-defined momenta. The *cross section*, which has the units of area, denoted by σ is the total number of events (of whatever type desired) divided by all of the quantities:

$$\sigma \equiv \frac{\text{Number of scattering events}}{\rho_A l_A \rho_B l_B A} \quad (2.36)$$

where \mathcal{A} are particles at rest with density ρ_A , aimed by particles of type \mathcal{B} with density ρ_B with velocity v , and l_A and l_B are the lengths of the bunches of particles. A is the cross-sectional

area common to the two bunches. We get

$$\text{Number of events} = \sigma l_{\mathcal{A}} l_{\mathcal{B}} \int d^2x \rho_{\mathcal{A}}(x) \rho_{\mathcal{B}}(x) \quad (2.37)$$

The *differential cross section* is $d\sigma/(d^3p_1 \dots d^3p_n)$ which when integrated over any small $d^3p_1 \dots d^3p_n$ gives the cross section for scattering into that region of final-state momentum space. Cross sections are computed for a production of a specific process. The *decay rate* Γ of an unstable particle \mathcal{A} assumed to be at rest into a specified final state is defined as

$$\Gamma \equiv \frac{\text{Number of decays per unit time}}{\text{Number of } \mathcal{A} \text{ particles present}}. \quad (2.38)$$

2.3.4.2 Interacting Fields

The example that was used in previous sections, the Klein-Gordon field, was a free field theory. No interactions and no scattering were involved. In reality, particles do interact and scatter off each other. In order to obtain such interactions, nonlinear terms must be added to the Lagrangian. One example of such interacting Lagrangian is the "phi-fourth" theory,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \quad (2.39)$$

where λ is a dimensionless *coupling constant*. The goal is to be able to compute scattering amplitude for an interacting theory, in order to convert them into cross sections. This is generally impossible to solve exactly. Instead it is computed in the framework of *perturbation theory*. It turns out that the perturbation series is quite simple in structure, and can be visualized with the use of *Feynman diagrams*.

2.3.5 Feynman Diagrams

In order to compute cross sections and decay rates, one must compute matrix elements of the S-matrix. The S-matrix gives the probability amplitude for a scattering event between *in* and *out* states. The probability amplitude for producing the final state is simply related to the cross section. Computing the S-matrix elements, or scattering amplitudes, is done differently depending on the quantization scheme, canonical or path integrals. As previously mentioned, the computation is done in a perturbation series. A Feynman diagram is a graphical representation of a perturbative contribution to the transition amplitude or correlation function. In the canonical quantization, a Feynman diagram represents a term in the Wick's expansion of the perturbative S-matrix.

The quantization scheme also provides *Feynman rules* in order to compute the value of each Feynman diagram. It involves providing a mathematical expression for a *propagator* for each internal line, or a virtual particle. Each diagram then has an amplitude which is a term in the perturbative expansion. In Figure 2.1, an example of a tree-level diagram representing a process of $e^+e^- \rightarrow \mu^+\mu^-$ through a photon γ is shown.

2.3.6 Renormalization

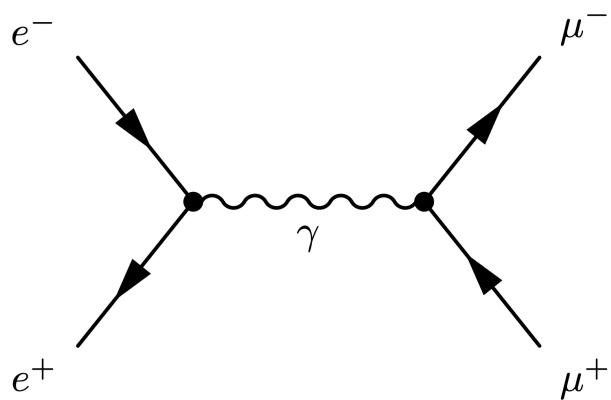


Figure 2.1: Feynman diagram representing the tree-level process of $e^+e^- \rightarrow \mu^+\mu^-$. Electron and positron annihilate each other and produce muon-antimuon pair through a virtual photon.

2.4 Symmetries

In 1964, Richard Feynman gave a series of seven lectures called *The Character of Physical Law* at Cornell University, as part of the Messenger Lectures series. The lectures were videotaped by the BBC, and are available online alongside their transcripts [7]. It is such a precious piece of history, and I would recommend everyone to watch it, as it is meant for the public audience. It showcases not only Feynman's ability to explain complex ideas and theories, but also his funny and enchanting personality. The fourth lecture in that series is called *Symmetry in Physical Law*, in which he starts by describing how Weyl defines a symmetry:

So, Weyl said, a thing is symmetrical if there's something that you can do to it, so that after you're finished doing it, it looks the same as it did before. That is the sense in which we say that the laws of physics are symmetrical: that there are things that we can do to the physical laws, or to our way of representing the physical laws, which make no difference and leave everything unchanged in its effects.

Most physicists will probably agree that that's what they mean when they say that a physical law is symmetrical. In this chapter we are concerned with symmetries of laws, rather than symmetries of objects, such as a human face, for example. In modern physics, hugely thanks to Noether's theorem, symmetries became fundamental, and they set the foundations of the Standard Model of particle physics. In this chapter I summarize the reasons why symmetries are so important, and what roles they play in the Standard Model. But first, I would like to introduce my favorite physics riddle, which Feynman described, starting at minute 40:04 in the same lecture:

Suppose that we were in telephone conversation with a Martian or an Arcturian, or something. We don't know where he is and we would like to describe things to him. We want to tell him about things. You say, so how's he going to understand the words, well, that's been studied very much by Professor Morrison here. He has pointed out that one way would be to start out and say tick tick two, tick tick tick three, and so on, and pretty soon the guy'd catch onto the numbers. Then—as he understands your number system, then—you can write lots of numbers and you could, for example, write a whole sequence of numbers that represents the weights, the proportional weights, of the different atoms, in succession.

Then say, hydrogen: 1.008, deuterium, and so on and so on. And he would—after he sat down with all those numbers and piddled around a while, would—discover that the mathematical ratios were the same as the ratios of the weights of the elements and, therefore, those names must be for the elements—and so gradually, you could, in talking to him, have a common language, in many ways, common. There are many—now comes the problem.

Suppose that he says, you fellas—after we get familiar with him, he says, "You're very nice; now I'd like to know what you look like." And you start out, "Well, we're about six feet tall." He says, "Six feet, how big is a foot?" "It's very easy," you say; "six feet tall is 170 thousand million hydrogen atoms high." Well, it's not a joke; it's a possible way of describing six feet to someone that has no measure, assuming that we cannot send him any samples, nor can we both look at the same object.

If we have to tell him how big we are; we can do it. That's because the laws of physics are not unchanged under a scale change. We can use that factor, use the properties of the scale to determine—I mean, you can use that fact to determine the scale.

Well, here we've described ourselves after telling six feet tall, and we're so-and-so bilateral on the outside, and we look like this, and there are these prongs sticking out, and all this. And he says, "That's very interesting; what do you look like on the inside?" So we describe the heart and so on, and we say, "Now, put the heart in on the left side."

Now the question is, how can we tell him which side is the left side?

Putting it short, the riddle is how to communicate the concept of left and right via radio signals to an alien civilization without reference to common sightings? Most physicists attempt this problem by thinking of the right-hand rule in electromagnetism. Unfortunately, that would not work. The correct solution is very unexpected, and one of the most surprising results of modern physics. The solution is explain in the section about parity symmetry 2.4.4.1.

2.4.1 Conservation Laws

A conservation law states that a particular measurable property of an isolated physical system does not change as the system evolves over time. Conservation laws are useful because they describe which processes can or cannot occur in nature. They also allow properties of the motion to be derived without solving the equations of motion. Conservation laws have changed throughout history, and some are a part of one theory but not another. Mass, for example, is conserved in classical mechanics, but not in relativity due to the principle of mass–energy equivalence. However, it a good approximation when the assumptions of classical mechanics hold, such as low velocities and energies, and large enough objects. Most conservation laws are exact, or absolute, in the sense that they apply to all possible processes. Some conservation laws are partial, in that they hold for some processes but not for others.

In continuum mechanics, the most general form of an exact conservation law is given by a continuity equation. For example, conservation of electric charge q is

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{j}, \quad (2.40)$$

where ρ is the density of q (amount per unit volume), \mathbf{j} is the flux of q (amount crossing a unit area in unit time), and t is time. There are several methods for identifying conservation laws. It can be hypothesize and proved mathematically. In classical mechanics, Hamilton–Jacobi equations provide a method for identifying constants of motion, and so does Poisson's theorem. In quantum mechanics, an observable quantity Q will be a constant of motion if it commutes with the hamiltonian, and it does not itself depend explicitly on time. In the context of particle physics, the most powerful theorem invoked in order to study conservation laws is Noether's theorem 2.4.2. It also this theorem that connects conservation laws to symmetries, which is the topic of this section.

Currently, exact conservation laws that have never been proven to be violated include conservation of mass-energy E , conservation of linear momentum \mathbf{p} , conservation of angular momentum \mathbf{L} , conservation of electric charge, conservation of weak isospin, conservation of color charge, conservation of boost 3-vector, and conservation of CPT parity. Approximate conservation laws, i.e., conservation laws which are approximately true in particular situations, such as low speeds, short time scales, or certain interactions include conservation of mechanical energy, mass, baryon number, lepton number, flavor, strangeness, space-parity, charge-parity, time-parity, and CP parity.

2.4.2 Noether's Theorem

In 1915, German-Jewish female mathematician Amalie Emmy Noether proved one of the most fundamental theorems of 20th century's modern physics. In spring of that year, she was invited by David Hilbert and Felix Klein to join the Göttingen mathematics department, challenging the views of some of his colleagues that a woman should not be allowed to teach at a university. Soon after arriving at Göttingen, she demonstrated her capabilities by proving the theorem now known as Noether's theorem, which shows that a conservation law is associated with any differentiable symmetry of a physical system [8]. That made modern theoretical physicists much more focused on symmetries.

Informally, Noether's theorem can be stated as:

If a system has a continuous symmetry property, then there are corresponding quantities whose values are conserved in time.

In the context of classical field theory, it can be stated as:

To each differentiable symmetry of a local Lagrangian, there corresponds a conserved current.

Previously, we described a symmetry by something that you can do, so that it looks the same as it did before. More formally in this context, a symmetry is the covariance of the form that a physical law takes, where by covariance, we mean invariance of the form of physical laws under differentiable transformations. That means, continuous transformation that leave the equations of motion invariant invariant. Formally, in the context of classical field theories, the theorem can be stated in the language of fields [1]. Consider a continuous transformations on the fields ϕ , which in infinitesimal form can be written

$$\phi(x) \rightarrow \phi'(x) = \phi(x) + \alpha\Delta\phi(x), \quad (2.41)$$

where α is an infinitesimal parameter and $\Delta\phi(x)$ is some deformation of the field configuration. It is considered a symmetry if it leaves the equations of motion invariant. This is insured either if the action is invariant under the transformation, or it is changed by a surface term. The Lagrangian, therefore, must be invariant under 2.41 up to a 4-divergence:

$$\mathcal{L}(x) \rightarrow \mathcal{L}(x) + \alpha\partial_\mu\mathcal{J}^\mu(x), \quad (2.42)$$

for some \mathcal{J}^μ . After varying the fields, one finds:

$$\partial_\mu j^\mu, \quad \text{for } j^\mu(x) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi - \mathcal{J}^\mu. \quad (2.43)$$

This result states that the current $j^\mu(x)$ is conserved. This can also be expressed by saying that the charge

$$Q \equiv \int_{\text{all space}} j^0 d^3x \quad (2.44)$$

is a constant in time.

2.4.3 Groups

2.4.4 Symmetries of the Standard Model

2.4.4.1 Parity

2.5 The Standard Model

2.5.1 Shortcomings of the Standard Model

2.6 Supersymmetry

2.7 Motivational Discussion for Supersymmetry

2.7.1 No go theorem

2.7.2 Dark Matter Candidate

2.7.3 Naturalness

Chapter 3

Experimental setup

One of the most useful methods to study the subatomic world of particle physics uses particle colliders. In such machines, particles are accelerated to very high speeds and energies, and smashed into each other. The particles that emerge from the collisions are then measured in a particle detector and then studied and analyzed. At the time of writing this thesis, the world's largest and highest-energy collider to date is the Large Hadron Collider (LHC) located in Geneva, Switzerland, operated by the European Organization for Nuclear Research (CERN). For the present work, data from the Compact Muon Solenoid (CMS) experiment has been analyzed. In this chapter, the LHC is described in 3.1, while the CMS experiment is described in 3.2.

3.1 The Large Hadron Collider

The LHC [9–11] is a circular hadron collider located at CERN near Geneva. It has been built inside the tunnel of the former Large Electron-Positron Collider (LEP) and has a circumference of about 27 km. Its tunnel is located as deep as 175 metres beneath the France–Switzerland border. The LHC was designed to deliver proton-proton (pp) collisions at a center-of-mass energy of up to $\sqrt{s} = 14 \text{ TeV}$ and heavy ion (lead-lead) collisions of up to $\sqrt{s} = 5.5 \text{ TeV}$ per nucleon. During Run 2 of the LHC (2015–2018), the center-of-mass energy of the pp collisions was $\sqrt{s} = 13 \text{ TeV}$.

The collider has four crossing points where the accelerated particles collide, as can be seen in the illustration in Figure 3.1. Nine detectors have been constructed at the LHC, located underground in large caverns excavated at the LHC's intersection points. Two of them, the ATLAS experiment and CMS, are large general-purpose particle detectors. The other detectors have more specialized roles. ATLAS and CMS measure a variety of SM physics, such as Higgs boson and top quark measurements, as well as look for BSM physics.

Two of the most interesting parameters of a particle collider are the center-of-mass energy of the collisions, and the luminosity. Due to energy conservation, the higher the energy of the collision, the heavier a theoretical particle can be produced. Therefore, in order to probe more massive theoretical particles (such as DM candidates in SUSY), the higher collision energy is necessary. The second parameter is the luminosity. The instantaneous luminosity depends on machine parameters and is given by

$$\mathcal{L} = \frac{1}{\sigma} dN t, \tag{3.1}$$

where σ is the corresponding cross section and dN/dt is the rate of particle interactions. The

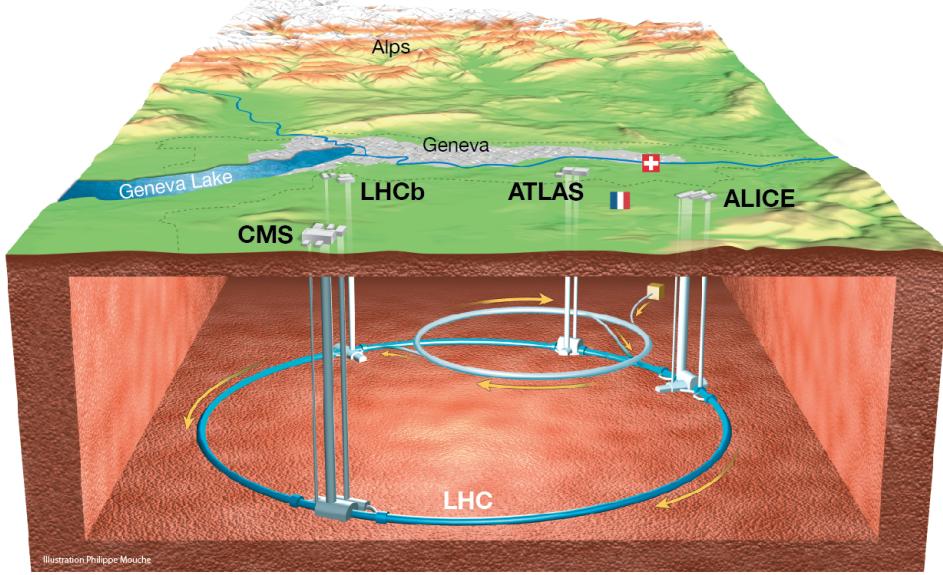


Figure 3.1: The LHC and the four main experiments located at the four interaction regions.

integrated luminosity is given by integrating the instantaneous luminosity over a period of time

$$L = \int \mathcal{L} dt. \quad (3.2)$$

The number of expected events N for a given process can be expressed in terms of the corresponding cross section σ times the integrated luminosity L

$$N = L \cdot \sigma. \quad (3.3)$$

Therefore, for rare processes, i.e., processes with very low cross section σ , access to large enough number of produced events requires higher integrated luminosity L . The integrated luminosity recorded in run 2 in CMS was around 138 fb^{-1} .

3.2 The Compact Muon Solenoid experiment

The Compact Muon Solenoid (CMS) experiment is one of two large general-purpose particle physics detectors built on the Large Hadron Collider (LHC) at CERN in Switzerland and France, as previously mentioned. The CMS apparatus has an overall length of 22 m, a diameter of 15 m, and weighs 14 000 tonnes. The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [12]. A cutaway diagram of the CMS detector can be seen in Figure 3.2.

Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μs [13]. The

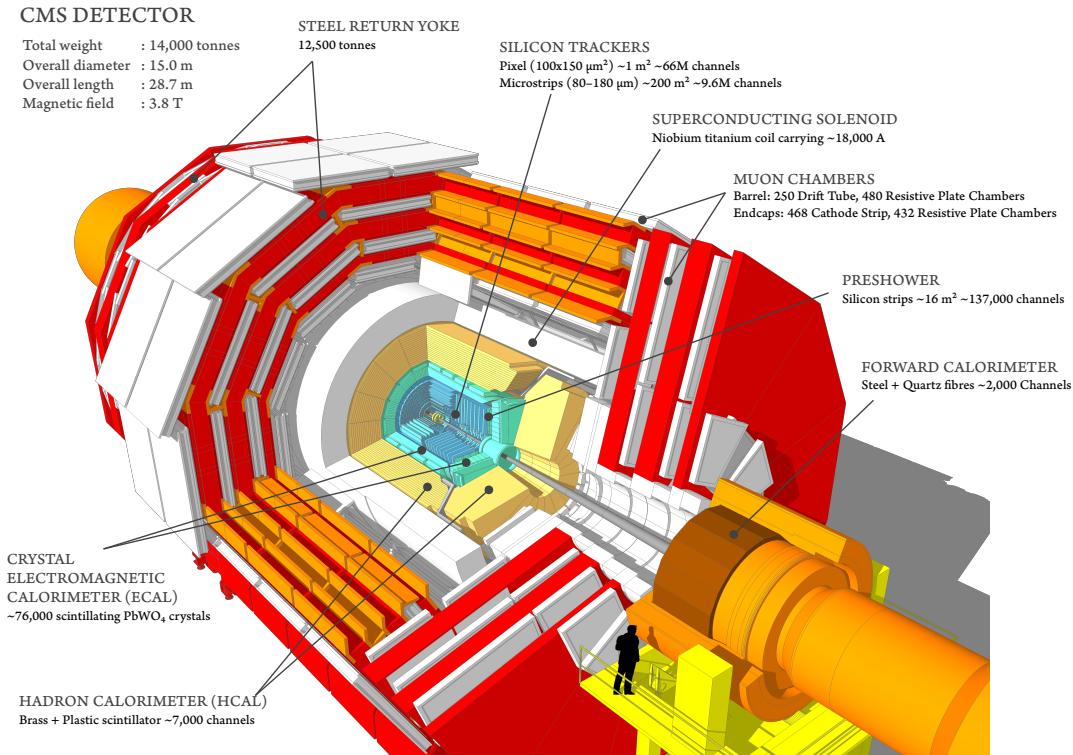


Figure 3.2: A cutaway diagram of the CMS detector.

second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [14].

The global event reconstruction (also called particle-flow event reconstruction [15]) aims to reconstruct and identify each individual particle in an event, with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. Photons (e.g. coming from π^0 decays or from electron bremsstrahlung) are identified as ECAL energy clusters not linked to the extrapolation of any charged particle trajectory to the ECAL. Electrons (e.g. coming from photon conversions in the tracker material or from B hadron semileptonic decays) are identified as a primary charged particle track and potentially many ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted along the way through the tracker material. Muons (e.g. from B hadron semileptonic decays) are identified as tracks in the central tracker consistent with either a track or several hits in the muon system, and associated with calorimeter deposits compatible with the muon hypothesis. Charged hadrons are identified as charged particle tracks neither identified as electrons, nor as muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as a combined ECAL and HCAL energy excess with respect to the expected charged hadron energy deposit.

The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the track momentum at the main interaction vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons attached to the track. The energy of muons is obtained from the corresponding track momentum. The energy of charged hadrons is determined from a combination of the track momentum and the corresponding ECAL and HCAL energies, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding

corrected ECAL and HCAL energies.

Jets are reconstructed offline from the energy deposits in the calorimeter towers, clustered using the anti- k_T algorithm [16, 17] with a distance parameter of 0.4. In this process, the contribution from each calorimeter tower is assigned a momentum, the absolute value and the direction of which are given by the energy measured in the tower, and the coordinates of the tower. The raw jet energy is obtained from the sum of the tower energies, and the raw jet momentum by the vectorial sum of the tower momenta, which results in a nonzero jet mass. The raw jet energies are then corrected to establish a relative uniform response of the calorimeter in η and a calibrated absolute response in transverse momentum p_T .

The primary vertex (PV) is taken to be the vertex corresponding to the hardest scattering in the event, evaluated using tracking information alone, as described in Section 9.4.1 of Ref. [18]. The missing transverse momentum vector \vec{p}_T^{miss} is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as p_T^{miss} [19]. The \vec{p}_T^{miss} is modified to account for corrections to the energy scale of the reconstructed jets in the event.

The silicon tracker used in 2016 measured charged particles within the range $|\eta| < 2.5$. For nonisolated particles of $1 < p_T < 10 \text{ GeV}$ and $|\eta| < 1.4$, the track resolutions were typically 1.5% in p_T and 25–90 (45–150) μm in the transverse (longitudinal) impact parameter [20]. At the start of 2017, a new pixel detector was installed [21]; the upgraded tracker measured particles up to $|\eta| < 3.0$ with typical resolutions of 1.5% in p_T and 20–75 μm in the transverse impact parameter [22] for nonisolated particles of $1 < p_T < 10 \text{ GeV}$.

Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. The single muon trigger efficiency exceeds 90% over the full η range, and the efficiency to reconstruct and identify muons is greater than 96%. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution, for muons with p_T up to 100 GeV, of 1% in the barrel and 3% in the endcaps. The p_T resolution in the barrel is better than 7% for muons with p_T up to 1 TeV [23].

The integrated luminosities for the 2016, 2017, and 2018 data-taking years have 1.2–2.5% individual uncertainties [24–26], while the overall uncertainty for the 2016–2018 period is 1.6%.

Chapter 4

Search for compressed Higgsinos with soft lepton tracks

4.1 Signal models

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

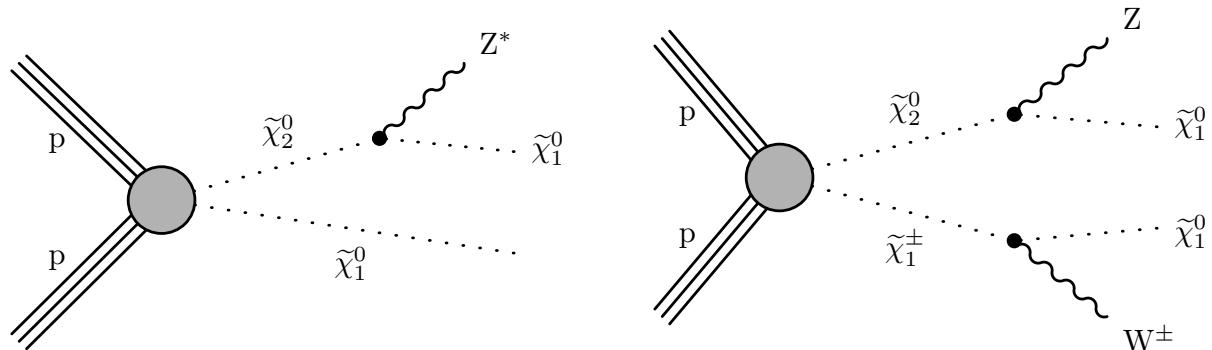


Figure 4.1: Production and decay of electroweakinos in the higgsino simplified model through $\tilde{\chi}_2^0 \tilde{\chi}_1^0$ (left) and $\tilde{\chi}_2^0 \tilde{\chi}_1^\pm$ (right).

Fixme Note: make sure we define the different deltaM somewhere

4.2 Previous searches

4.3 Search strategy

4.3.1 Final state with two muons

4.3.2 Final state with one lepton and one track

4.4 Signal signature and base selection

To develop an effective analysis strategy, the signal kinematics are studied and exploited. The production and decay of electroweakinos give rise to unique event characteristics that can be leveraged to differentiate the signal from the Standard Model (SM) background. Distributions of key observables from signal and background processes are compared in order to define a pre-selection or set of base cuts that retains the maximum signal while rejecting as much background as possible. All the following distributions of key observables were generated by weighting the simulated data to the 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 $|\eta| < 2.4$. Discussion is provided for each event property, and the incorporation of additional selection criteria is specified building toward the base selection.

4.4.1 Missing transverse energy

A driving factor for most searches for Dark Matter (DM) at the LHC is the presence of a DM candidate in the final state. The identity and properties of the particle (or particles in the case of multiple DM candidates) vary, but they do have much in common. In this Supersymmetry (SUSY) search, the DM candidate is the Lightest Supersymmetric Particle (LSP), assumed to be a neutralino. A neutral particle that does not interact electromagnetically or via the strong force (i.e., is colorless) will not be detected and will leave traces in the form of a transverse momentum imbalance, which is referred to as E_T^{miss} (missing transverse energy or missing transverse momentum). Because of R-parity conservation, the signal contains two DM candidates in the final state, which are the LSPs, the neutralinos $\tilde{\chi}_1^0$. Therefore, a considerable magnitude of E_T^{miss} is expected in the signal. As described in Section 4.6.4, a suitable proxy for the E_T^{miss} is the missing transverse hadronic energy, or H_T^{miss} , which is highly correlated with E_T^{miss} , but better suited to the definition of lepton isolation and its use in the background estimation methods. Both E_T^{miss} and H_T^{miss} observables are examined in Figure 4.2.

As expected, E_T^{miss} and H_T^{miss} are largely unaffected by the different choices for Δm , while the higgsino parameter μ affects the distributions mainly through its falling production cross section as a function of the higgsino parameter μ . The region of interest in order to be efficient with respect to the triggers is located at $H_T^{\text{miss}} \geq 220 \text{ GeV}$, as discussed in Section 4.7. Although this is a harsh and inefficient cut, it becomes apparent when examining the SM background in both regions of $H_T^{\text{miss}} < 220 \text{ GeV}$ and $H_T^{\text{miss}} \geq 220 \text{ GeV}$ to conclude that most of the sensitivity comes from the $H_T^{\text{miss}} \geq 220 \text{ GeV}$ region, as the production of real H_T^{miss} (or E_T^{miss}) results from the production of neutrinos in the event. These are much less common than Quantum Chromodynamics (QCD) events that dominate the $H_T^{\text{miss}} < 220 \text{ GeV}$ region.

4.4.2 Jets and hardronic activity

As mentioned in the previous section, signal events tend to have small momentum imbalance. In order to induce significant missing transverse energy, some additional activity must take place within the events, and this most often comes in the form of one or more Initial State Radiation (ISR) jets. An ISR jet is created when one of the incoming protons emits radiation (such as a quark or a gluon) before the interaction. If a jet with sufficiently high p_T is emitted, the remainder of the interaction is recoiled against this jet and imparts momentum onto the system of invisible particles in the opposite direction. As a result, the boosted neutralinos $\tilde{\chi}_1^0$ give rise to higher H_T^{miss} . As described in Section 4.6.5, the jets are required to have $p_T \geq 30 \text{ GeV}$ and be located within the tracker acceptance ($|\eta| < 2.4$). At least one such jet is required in each event. The distributions of the number of jets and the leading jet p_T are displayed in Figure 4.3.

The signal signature rarely includes a b-jet, that is, a jet resulting from the hadronization of a bottom quark. However, standard model top quark pair production leads to a large numbers

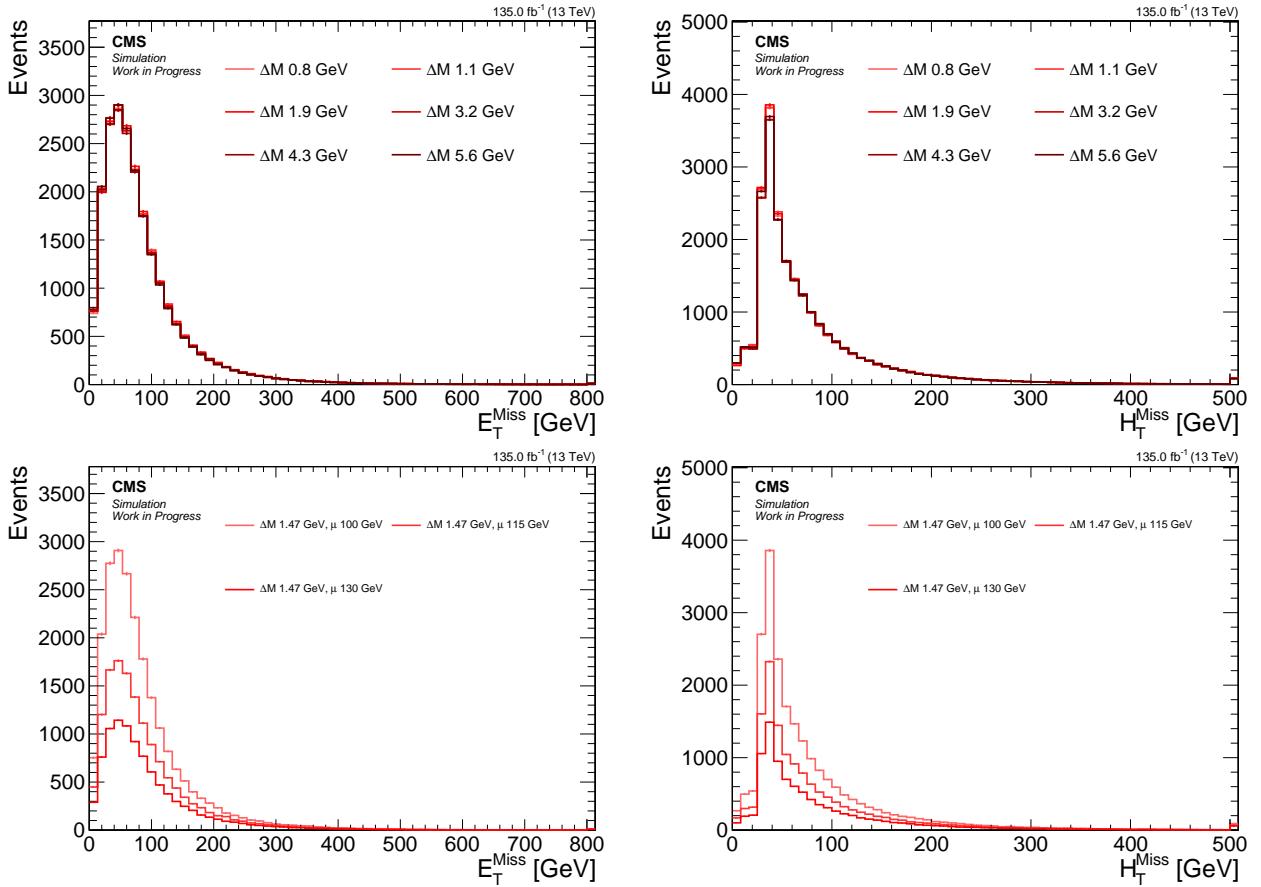


Figure 4.2: Signal distributions of E_T^{miss} (left) and H_T^{miss} (right) comparing various Δm with a fixed higgsino parameter $\mu = 100\text{ GeV}$ (upper), and comparing various μ with fixed $\Delta m = 1.47\text{ GeV}$ (lower).

of events with significant missing transverse energy and two or more b-jets. To reject this background, events are vetoed if a b-jet is identified in the event. As described in Section 4.6.5, the DEEPCSV bottom flavor tagging discriminant with a medium working point is used. The multiplicity of b-tagged jets is shown in Figure 4.4, where the choice of number of b-tagged jets equals to zero appears well-justified.

As an ISR jet is required in the event, it is expected that the E_T^{miss} and the H_T^{miss} will be directed in the opposite direction of the jet, or at an azimuthal angle close to π . This feature is not as clearly observed in events with multiple jets in the SM background, such as those arising from QCD, where the missing transverse energy tends to align with the leading or sub-leading jet. To reduce the QCD background, a requirement of $\min \Delta\phi(\vec{H}_T^{\text{miss}}, \vec{j}) > 0.4$ is imposed.

4.4.3 Base selection

The section is recapped by summarizing the base selection of the analysis. The base selection, also known interchangeably as the preselection, is applied to all analysis categories. It is listed in Table 4.4.3.

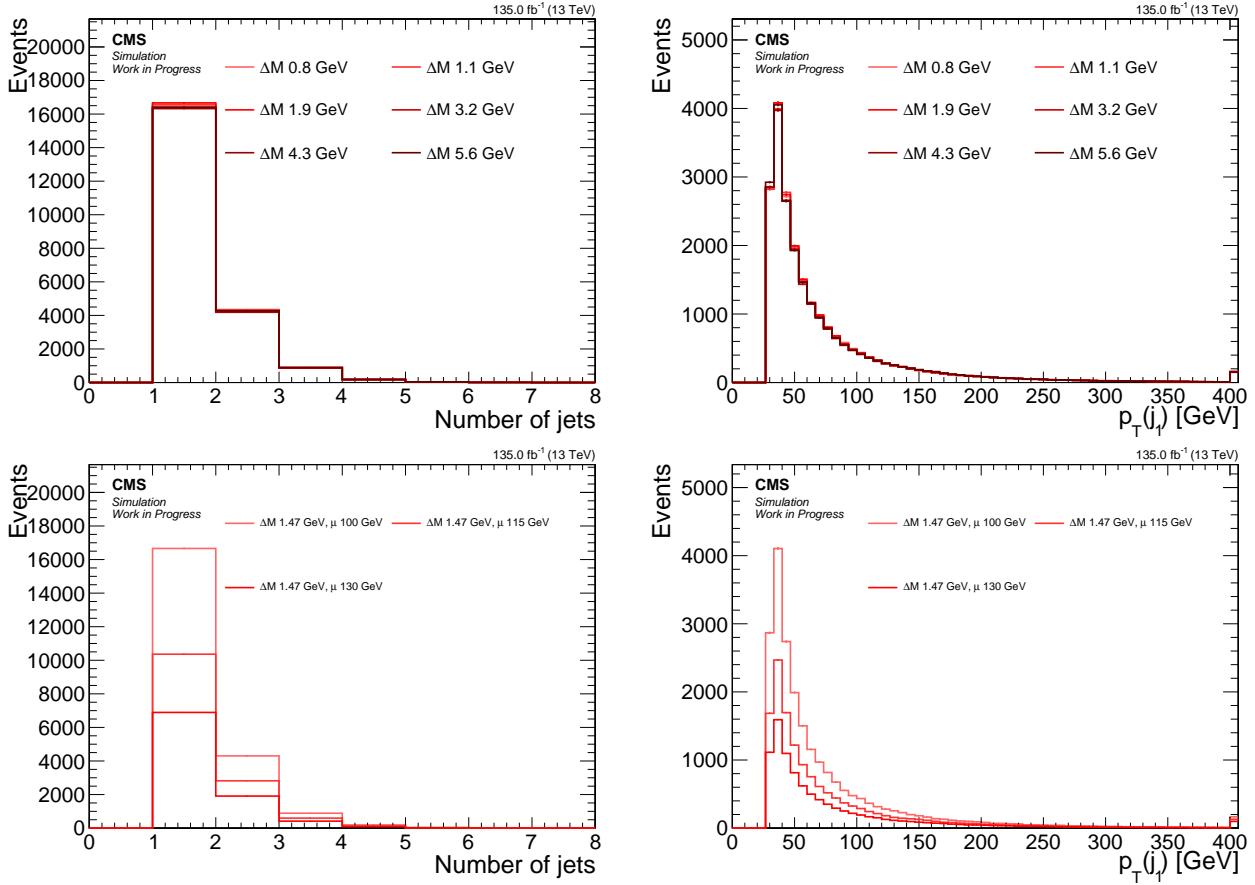


Figure 4.3: Signal distributions of *number of jets* (left) and *leading jet p_T* (right) comparing various Δm with a fixed higgsino parameter $\mu = 100$ GeV (upper), and comparing various μ with fixed $\Delta m = 1.47$ GeV (lower).

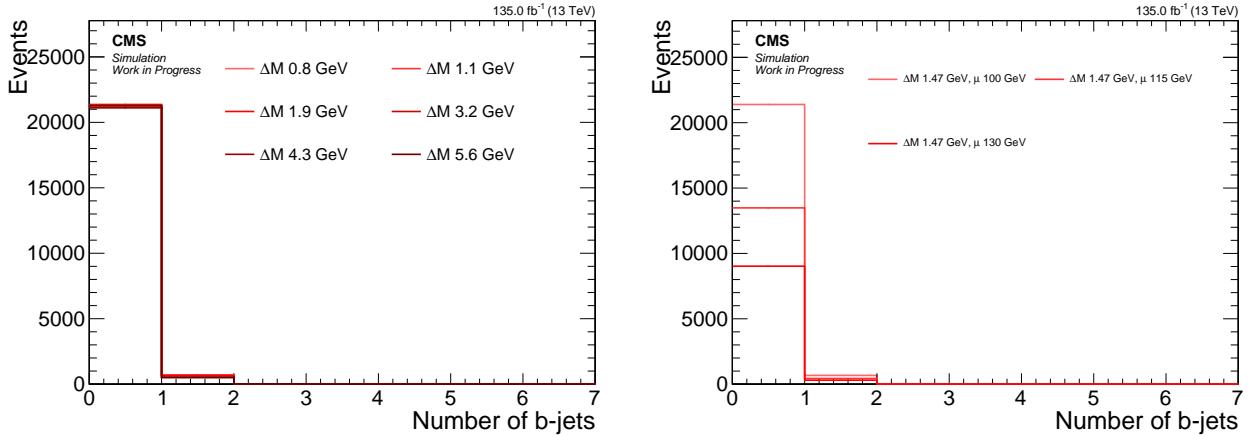


Figure 4.4: Signal distributions of *number of b-tagged jets* comparing various Δm with a fixed higgsino parameter $\mu = 100$ GeV (left), and comparing various μ with fixed $\Delta m = 1.47$ GeV (right).

Table 4.1: The preselection criteria, which are applied to all analysis categories.

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

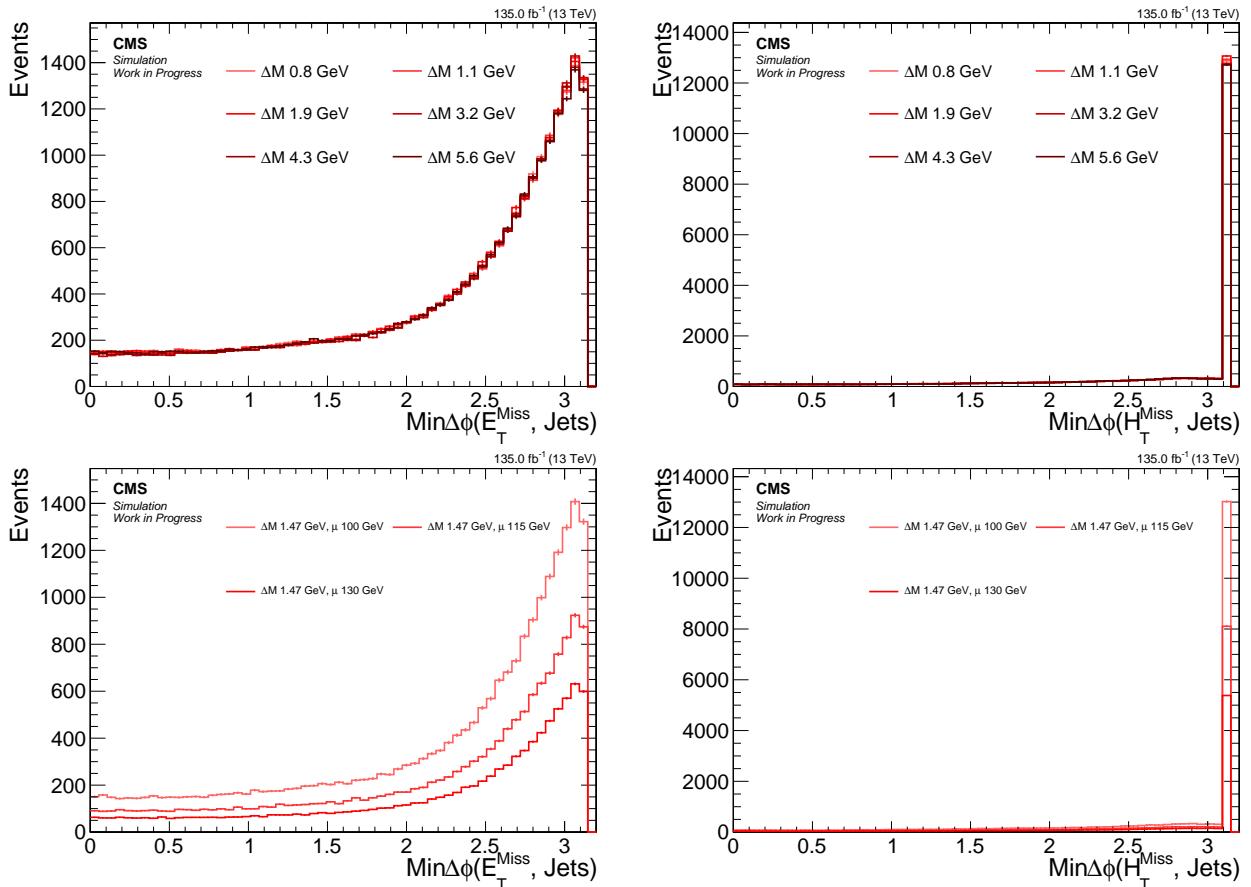


Figure 4.5: Signal distributions of $\text{min} \Delta\phi(\vec{E}_T^{\text{miss}}, \vec{j})$ (left) and $\text{min} \Delta\phi(\vec{H}_T^{\text{miss}}, \vec{j})$ (right) comparing various Δm with a fixed higgsino parameter $\mu = 100 \text{ GeV}$ (upper), and comparing various μ with fixed $\Delta m = 1.47 \text{ GeV}$ (lower).

4.4.4 Lepton kinematics

The hadronic component of signal events has been the focus up until this point. However, the dilepton system contains the most distinctive features of the signal. To fully understand the unique phase space of the dilepton system, generator level distributions are examined first, followed by an exploration of the effects of reconstruction on those observables. Since the dimuon category is the most sensitive and because the logic applies analogously to the two-electron final state, the electron category is excluded from the following sections. The lepton kinematics change dramatically as a function of Δm . In contrast, the higgsino parameter μ effects almost only the overall normalization due to the different production cross section. Therefore, the higgsino parameter is set to $\mu = 100$ GeV in the following sections, with the Δm varied.

4.4.4.1 Lepton η and transverse momentum p_T

The signal acceptance and sensitivity are significantly impacted by the thresholds of the transverse momentum p_T distribution of the muons that make it through the reconstruction and identification. The selection applied to the muons in this analysis is described in Section 4.6.2 and referred to as the *analysis selection*. This section aims to examine the importance of the p_T on the signal and its dilepton kinematic distributions.

The generator level distribution of p_T , or the so-called *truth* distributions, which do not exhibit any detector or reconstruction features, are examined first. The distribution of reconstructed p_T is then compared with the generator level distribution in Figure 4.6.

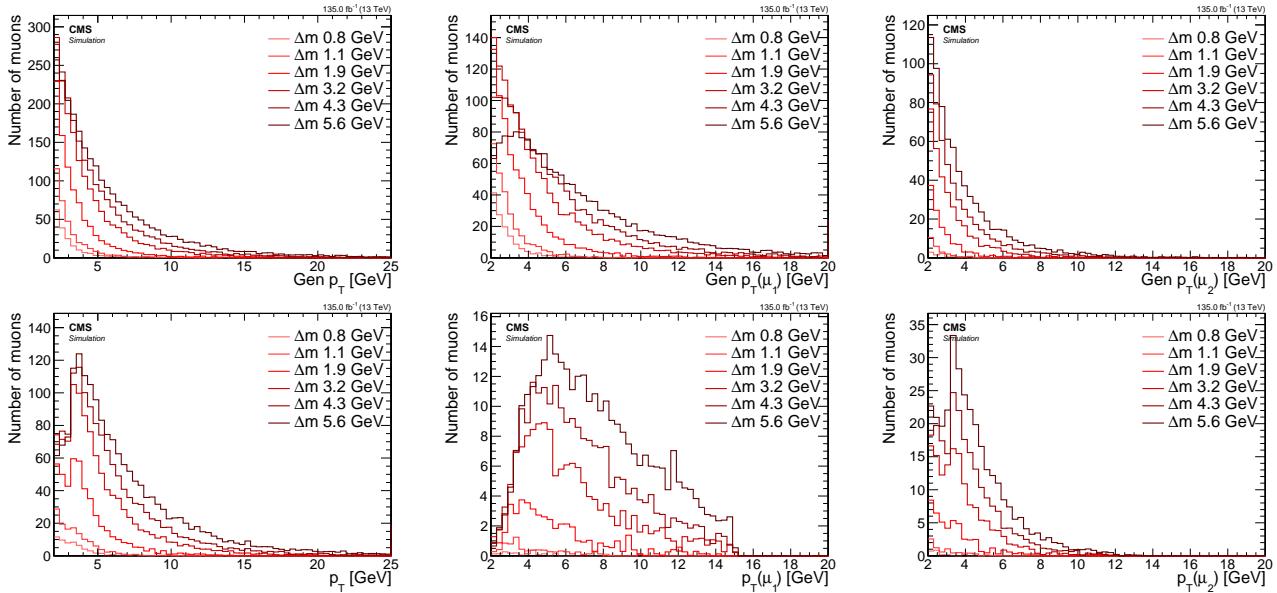


Figure 4.6: Signal p_T distributions for inclusive (left), leading muon μ_1 (middle), subleading muon μ_2 (right) at generator level (top) and reconstruction level passing analysis selection (bottom).

When comparing the generator level and reconstruction level inclusive p_T distributions, it becomes apparent that a reshaping occurs around 3 GeV. A significant proportion of the generated muons with $p_T < 3$ GeV are lost in the reconstruction process. The subleading muon p_T distribution at the reconstruction level has a camel shape, whereby the efficiency drops below a p_T of 3 GeV to about half its maximum value and is only partially regained at $p_T > 3$ GeV. This effect is due to the detector geometry and is more clearly visible when splitting the p_T distribution into a barrel ($|\eta| < 1.2$) and encaps ($|\eta| \geq 1.2$) portions, as shown in Figure 4.7.

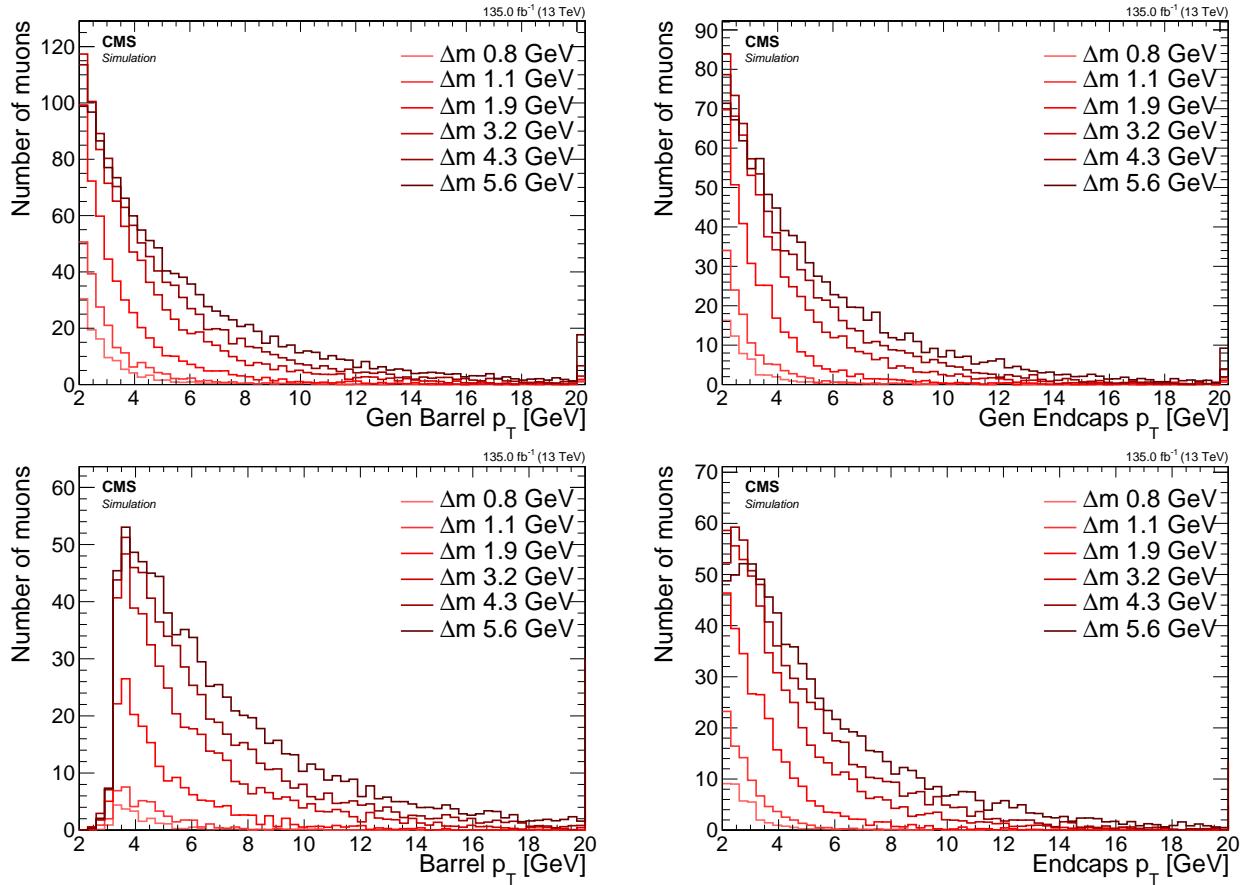


Figure 4.7: Signal inclusive p_T distributions for barrel $|\eta| < 1.2$ (left) and endcaps $|\eta| \geq 1.2$ (right) at generator level (top) and reconstruction level passing analysis selection (bottom).

When comparing the generator level distribution of the barrel muons on the top left with its reconstructed counterpart on the bottom left, Figure 4.7 shows that the barrel, shown on the left, is almost completely unable to reconstruct muons with $p_T < 3$ GeV, while the endcaps, shown on the right, are able to do so. As demonstrated in the upcoming sections on $m_{\ell\ell}$ and ΔR (see 4.4.4.2 and 4.4.4.3), the relationship between these observables has consequences for the reshaping of kinematic distributions, as well as for signal acceptance in general. Access to low Δm signal points is crucially dependent on the low p_T region of $2 \leq p_T \leq 3.5$ GeV, which is mainly achieved with the help of the muon chamber endcaps, as can be seen here.

Since the barrel and endcaps are separated by different regions of η , $|\eta| < 1.2$ for barrel and $|\eta| \geq 1.2$ for endcaps, the muon η distributions merit further examination as well. They can be seen at Figure 4.8. The dimuon analysis channel only selects muons within the tracker range of $|\eta| < 2.4$. This is why the muons with $|\eta| > 2.4$ are not present in the reconstruction plots on the bottom. It can be seen that the main effect of going from the inclusive $|\eta|$ at the 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, an examination of other kinematic variables of the dilepton system is now possible.

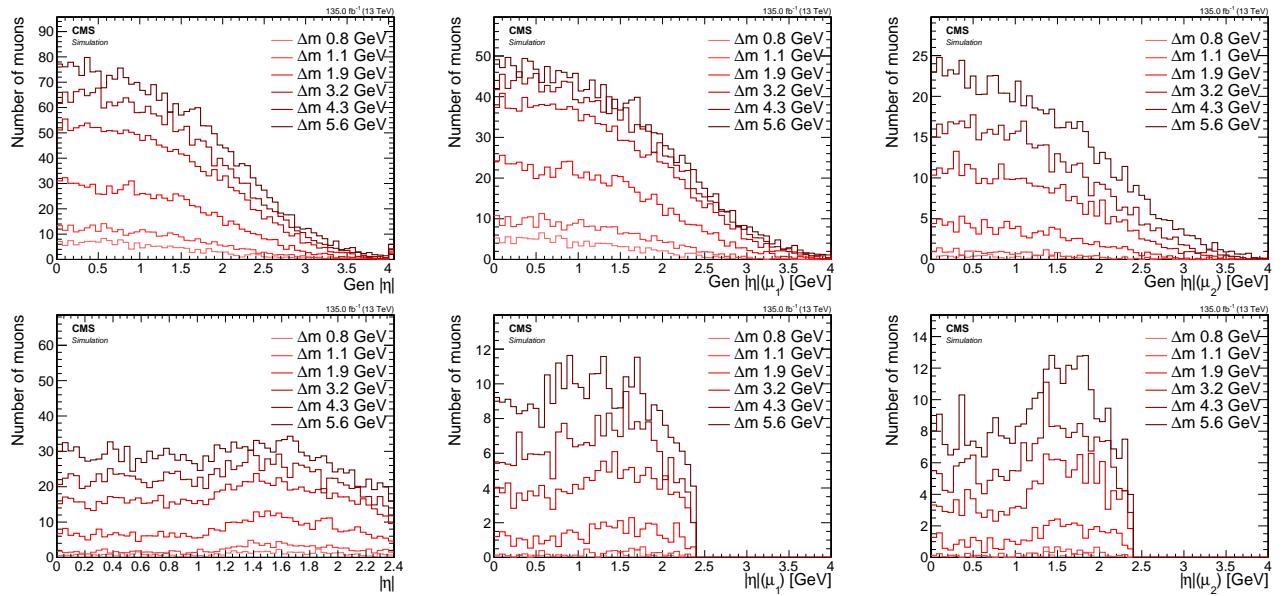


Figure 4.8: Signal $|\eta|$ distributions for inclusive (left), leading muon μ_1 (middle), subleading muon μ_2 (right) at generator level (top) and reconstruction level passing analysis selection (bottom).

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

The invariant mass of the two leptons resulting from the decay of the $\tilde{\chi}_2^0$ has a unique shape due to the limited allowed phase space of the 3-body decay. As 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 $\tilde{\chi}_1^0$, that is, Δm . Therefore, the $m_{\ell\ell}$ distribution is expected to have an edge at Δm . Distributions of the generator level invariant mass can be seen in Figure 4.9.

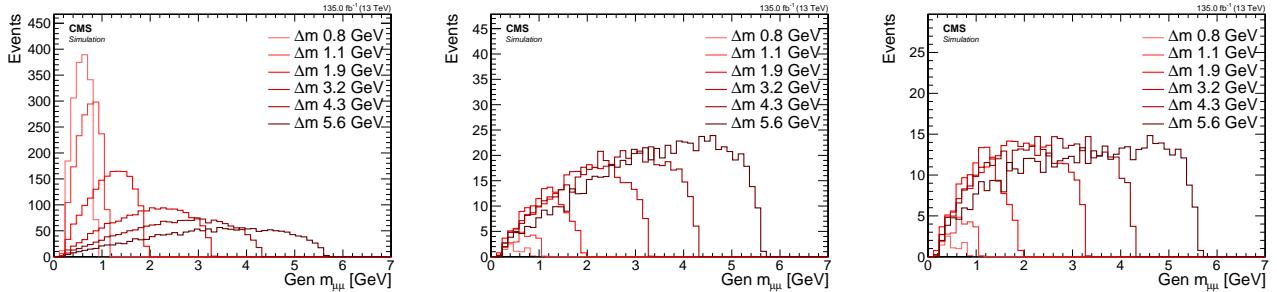


Figure 4.9: Signal generator level $m_{\ell\ell}$ distributions with no cuts (left), with $p_T(\mu_i) > 2 \text{ GeV}$, $i = 1, 2$ (middle) and with the SOS orthogonality condition: $p_T(\mu_i) > 2 \text{ GeV}$, $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R \leq 0.3$ (right).

The inclusive distribution of the invariant mass of the muons $m_{\mu\mu}$ is shown on the left. The edge of the $m_{\mu\mu}$ distribution for each signal point is located right at the corresponding Δm . However, when the muons p_T is required to be $p_T \geq 2 \text{ GeV}$, the shape of the distribution shifts, due to the lower efficiency for small Δm values, as depicted in the middle plot. Lastly, the effect of orthogonalizing phase space to the SOS analysis is demonstrated in the rightmost plot. The effect is strongest in high Δm and quite subtle in low Δm .

To explain the reshaping that occurs to the $m_{\mu\mu}$ distribution, the relationship between the p_T of the muons and the invariant mass is examined. One signal with low Δm of 1.13 GeV and one with high Δm of 5.63 GeV are selected for this analysis. The distributions are shown in Figure 4.10, leading muon denoted μ_1 while subleading muon is denoted μ_2 .

Earlier, it was established that the invariant mass distribution has an edge at Δm , and the value of Δm can be read from these plots. Another interesting feature is a lower edge in the Δm distribution at around $\sim 0.2 \text{ GeV}$, which is due to each muon having a mass of around $\sim 0.1 \text{ GeV}$. It is now clear that by requiring both muons to have $p_T \geq 2 \text{ GeV}$, a significant portion of the signal is lost. This effect becomes particularly substantial for the low $\Delta m = 1.13 \text{ GeV}$ (top row). The magnitude of this effect is quantified by a cutflow, shown in Table 4.4.4.2, 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 previous line. The first line the number of events with exactly 2 muons at the generator level with at least one jet with $p_T \geq 30 \text{ GeV}$ and $|\eta| < 2.4$. The event number is weighted to Run II luminosity of $\mathcal{L} = 135 \text{ fb}^{-1}$.

Table 4.2: Generator level efficiency on muons selections

Cut	Weighted 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	-	-
$p_T \geq 2 \text{ GeV}$	24.7	724.9	0.015	0.41
SOS orthogonality	24.7	490.6	1	0.68

Table 4.4.4.2 shows that for the low Δm of 1.13 GeV , the acceptance of the signal is significantly reduced by the $p_T \geq 2 \text{ GeV}$ cut, with only 1.5% of the signal remaining. In contrast, the

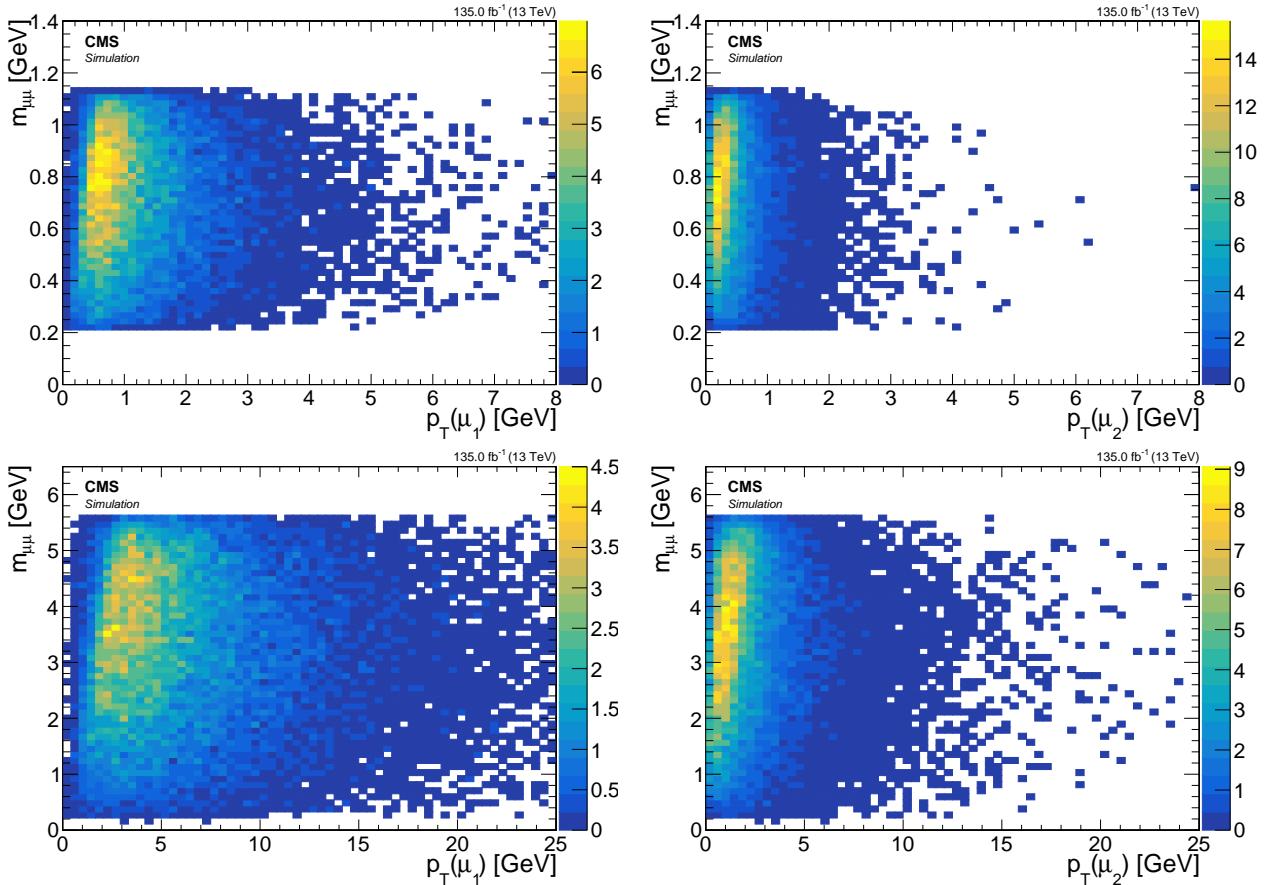


Figure 4.10: Signal $m_{\mu\mu}$ vs. p_T for leading lepton μ_1 (left) and subleading lepton μ_2 (right) for $\Delta m = 1.13 \text{ GeV}$ (top) and $\Delta m = 5.63 \text{ GeV}$ (bottom).

orthogonality condition of requiring $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R(\ell\ell) \leq 0.3$ does not affect it any further. The situation is different for the high Δm of 5.63 GeV , where the p_T cut rejects more than half of the signal and the SOS orthogonality condition rejects an additional two thirds.

It has been established that the p_T thresholds affect the $m_{\ell\ell}$ distribution due to the relationship between the two variables. Next, it is investigated how the reconstruction discussed in Section 4.4.4.1 impacts the $m_{\mu\mu}$ distribution. The distributions of the reconstructed $m_{\mu\mu}$ can be seen in Figure 4.11. Comparing these distributions to the two right plots in Figure 4.9 not only are fewer events surviving the reconstruction, but also some Δm model points are peaking between 1 GeV to 2 GeV with the SOS orthogonality condition applied.

4.4.4.3 Lepton separation ΔR

The lepton separation is defined by the equation $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, where η represents the pseudorapidity and ϕ is the azimuthal angle measured in radians. The value of ΔR is significant in this analysis because the produced leptons tend to be located in close proximity to each other and therefore are not easily isolated according to standard definitions. Special attention is given to ensuring that the collimated nature of the leptons can be used to differentiate signal leptons from the non-isolated leptons in the SM background. It is worth noting that, for the purposes of orthogonality, the requirement of $\Delta R(\ell\ell) > 0.3$ utilized in previous SOS analyses [27] is reverted.

Similar to the invariant mass discussed in Section 4.4.4.2, we examine the distributions of ΔR for various Δm options with different cuts applied to observe their effect. The left plot of Figure 4.12 shows that roughly the same number of events are produced for all Δm model

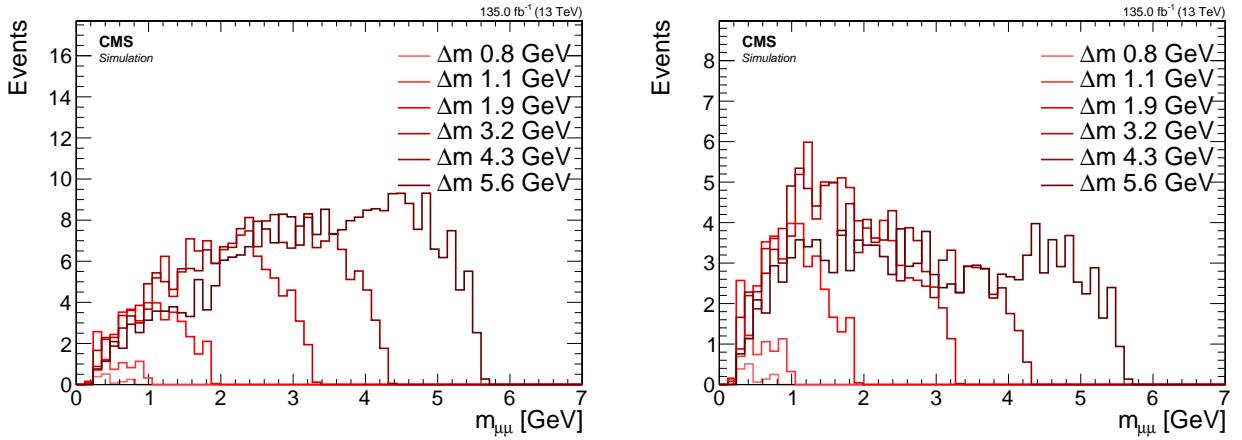


Figure 4.11: Distributions of reconstructed $m_{\mu\mu}$ in signal events with analysis selection (left) and the additional SOS orthogonality condition (right).

points. However, when applying a cut of $p_T(\mu) > 2 \text{ GeV}$, a hierarchy of Δm points emerges, with fewer 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 that must be satisfied.

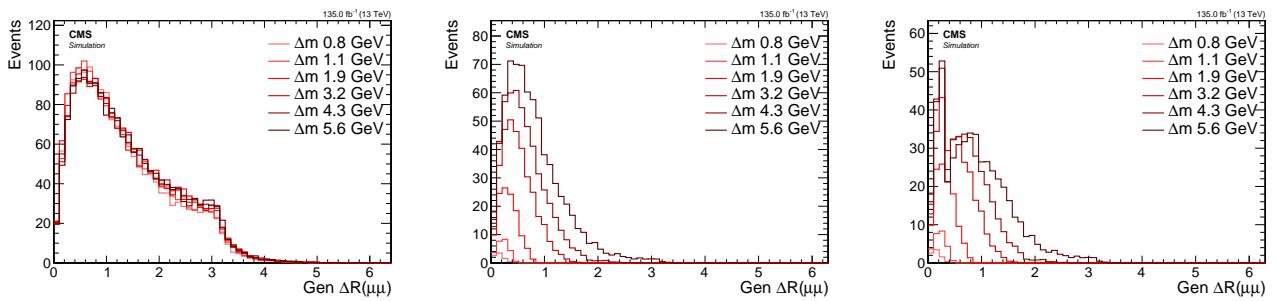


Figure 4.12: Signal generator level ΔR distributions with no cuts (left), with $p_T(\mu_i) > 2 \text{ GeV}$, $i = 1, 2$ (middle) and with SOS orthogonality condition $p_T(\mu_1) > 2 \text{ GeV}$, $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R \leq 0.3$ (right).

To understand the shaping and hierarchy formation due to the p_T cut, the p_T of the muons is plotted vs. $\Delta R(\ell\ell)$ in Figure 4.13. Requiring $p_T(\mu_2) \geq 2 \text{ GeV}$ for $\Delta m = 1.13 \text{ GeV}$ limits the range of $\Delta R(\mu\mu)$ to less than 0.4, while leaving a large range exceeding 3 for the $\Delta m = 5.63 \text{ GeV}$ model point. To gain access and sensitivity to the low Δm model points, allowing small $\Delta R(\ell\ell)$ values, less than 0.3 is necessary, even before considering the reconstruction efficiency of the leptons. In the next sections, the study of reconstructed leptons and the isolation criteria will enable the retention of signal points with highly-columnated lepton pairs, as further explored in Section 4.6.7.

As seen in Section 4.4.4.2 for $m_{\mu\mu}$, reconstruction has an effect on both the shape and overall count of events. The effects on the $\Delta R(\mu\mu)$ distributions are investigated in Figure 4.14.

Comparing Figure 4.14 and Figure 4.12, the main effect of the reconstruction on the $\Delta R(\mu\mu)$ is the overall normalization, which is due to reconstruction efficiency.

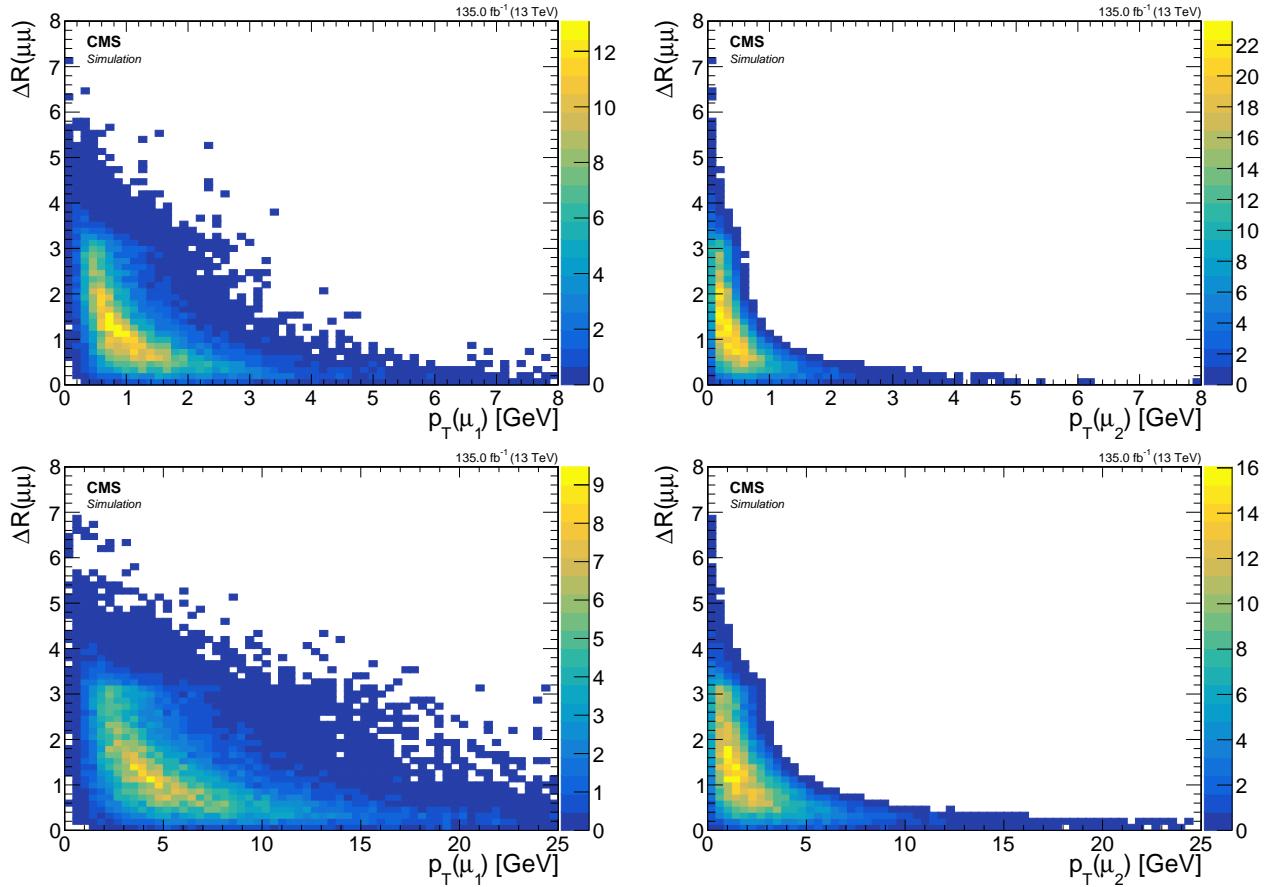


Figure 4.13: Event distributions in the plane of $\Delta R(\mu\mu)$ vs. p_T for leading lepton μ_1 (left) and subleading lepton μ_2 (right) for signal models with $\Delta m = 1.13 \text{ GeV}$ (top) and $\Delta m = 5.63 \text{ GeV}$ (bottom).

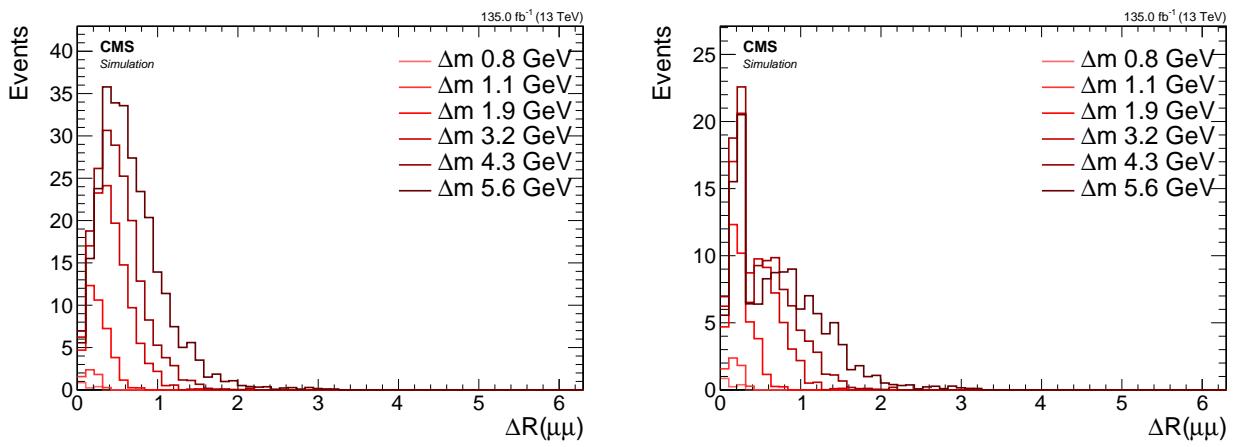


Figure 4.14: Distributions of the reconstructed $\Delta R(\mu\mu)$ with preselection applied (left) and the additional SOS orthogonality condition (right).

4.4.5 Main drivers of sensitivity

The above studies reveal the main drivers of the sensitivity to different model points of this analysis, and may inform future analysis strategies that expand on the current work. This section has not explicitly included SM background in the plots, making it hard to conclude what effects changing the cuts to E_T^{miss} or other event level observables might have. However, it is very clear from examining the dilepton kinematics that for low Δm model points, regions with low p_T and ΔR contain the bulk of the signal events. 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 .

The next sections will explore how to lower the threshold on the muon transverse momentum and deal with collimated leptons that might pose a challenge in regards to the isolation criterion.

4.5 Simulated samples

4.5.1 Standard Model simulated samples

4.5.2 Signal simulated samples

4.6 Object definition and selection

In Section ??, the reconstruction and identification of objects in the detector were presented. Additionally, the signal signature was studied in Section 4.4. In this section, a set of object selection criteria is devised to obtain a sample that is as pure as possible with respect to the signal leptons, while still retaining as much signal as possible. As discussed in Section 4.3, the focus is on selecting opposite-charge, same-flavor leptons $\ell^+\ell^-$ resulting from the $\tilde{\chi}_2^0$ that decays into a $\tilde{\chi}_1^0$ and a Z^* , i.e., $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0\ell^+\ell^-$. Two choices of Δm^0 are presented in the following section: a relatively high Δm^0 of $\Delta m^0 = 5.63$ GeV and a low Δm^0 of $\Delta m^0 = 1.92$ GeV, but not so low as to prevent enough electrons from surviving the initial reconstruction p_T threshold of 5 GeV. The higgsino parameter is fixed at $\mu = 100$ GeV.

In Section 4.4, the base selection required at least one jet in the event with $p_T \geq 30$ GeV and $|\eta| < 2.4$, without any other selection. However, unlike in that section, objects are not weighted to any luminosity in this section, as the focus is on the proportion between object types. Two types of reconstructed leptons are differentiated: those originating from the targeted decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0\ell^+\ell^-$, shown in blue, and those that do not, referred to as *other*, shown in yellow. Signal leptons are marked as such by matching a reconstructed lepton to a generator level lepton, which has been confirmed to have the $\tilde{\chi}_2^0$ as its parent. Leptons marked as *other* may have been misreconstructed, misidentified, or may be a result of the hadronisation process in a jet (such as the ISR jet). The goal is to select as many blue leptons as possible while rejecting as many yellow ones as possible.

In the following sections, the term *efficiency* refers to the proportion of signal leptons passing a selection, divided by the initial number of signal leptons, and the term *purity* refers to the proportion of signal leptons (blue) to the sum of the signal leptons and *other* leptons (yellow). The goal is to find selection criteria with high efficiency and high purity. However, these two quantities can sometimes compete with each other, requiring compromises.

4.6.1 Electrons

The electrons are subject to an initial lower threshold on the reconstructed p_T 5 GeV, and are reconstructed using a loose working point (WP), as described in Section ???. The first distribution of interest regarding the electrons is their angular separation from the leading jet in the event, denoted as $\Delta R(j_1, e)$. The distributions are shown in Figure 4.15. Two key features are apparent. The first has already been discussed in Section 4.4, which is that probing lower Δm necessitates access to low p_T leptons. The threshold of $p_T \geq 5$ GeV on the electrons leads to reduced signal acceptance. This is evident from the difference between the high and low Δm cases. The second interesting feature is that the signal electrons are predominantly located outside the leading jet. This is because the leading jet is typically an ISR jet, which boosts the $\tilde{\chi}_2^0\tilde{\chi}_1^0$ system away from it, causing them to be back-to-back. Thus, a cut of $\Delta R(j_1, e) > 0.4$ is made to account for this.

Distributions of electron p_T are examined by applying the $\Delta R(j_1, e) > 0.4$ cut. It is observed that the p_T distribution depends strongly on the Δm , as previously seen for muons in Section 4.4.4.1. Thus, a choice must be made regarding which Δm to prioritize, and the lower Δm case is chosen for increased sensitivity. However, the two choices are compared in Figure 4.16. As expected, the p_T distribution of the electrons falls more rapidly for the low Δm case. It is observed that there are very few electrons surviving above 15 GeV. Therefore, a cut of $p_T < 15$ GeV is chosen. The η distribution is seen in Figure 4.17, after the previous cuts to gain a better understanding of where most of the non-signal electrons originate from. For the $\Delta m = 1.92$ GeV case, it can be clearly seen that the endcaps of the electromagnetic calorimeter (ECAL) are performing worse compared to the barrel ($|\eta| < 1.48$). The transition is easily

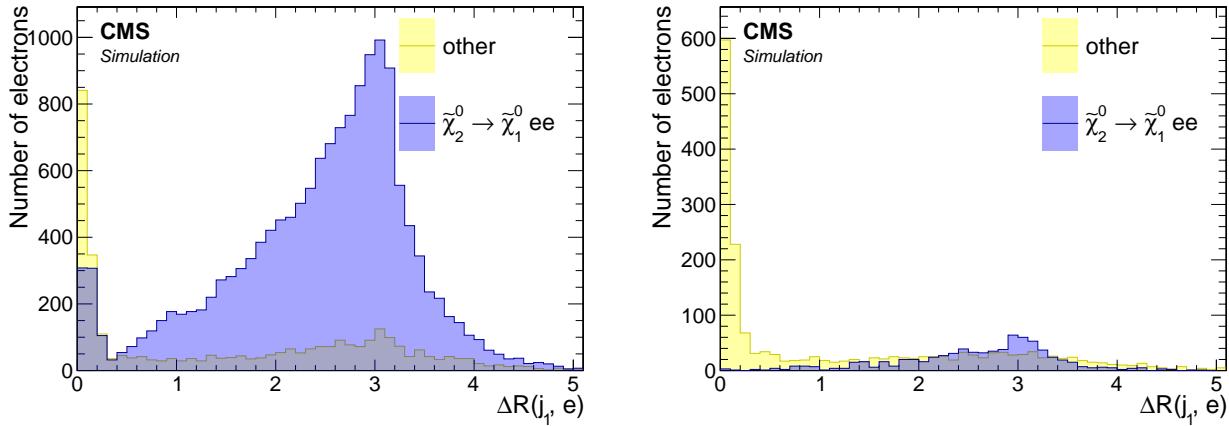


Figure 4.15: Angular separation between reconstructed electrons with loose ID and the leading jet $\Delta R(j_1, e)$ for $\Delta m = 5.63 \text{ GeV}$ (left) and $\Delta m = 1.92 \text{ GeV}$ (right).

noticeable through a sharp drop in purity at the transition. This effect is most pronounced for low- p_T electrons.

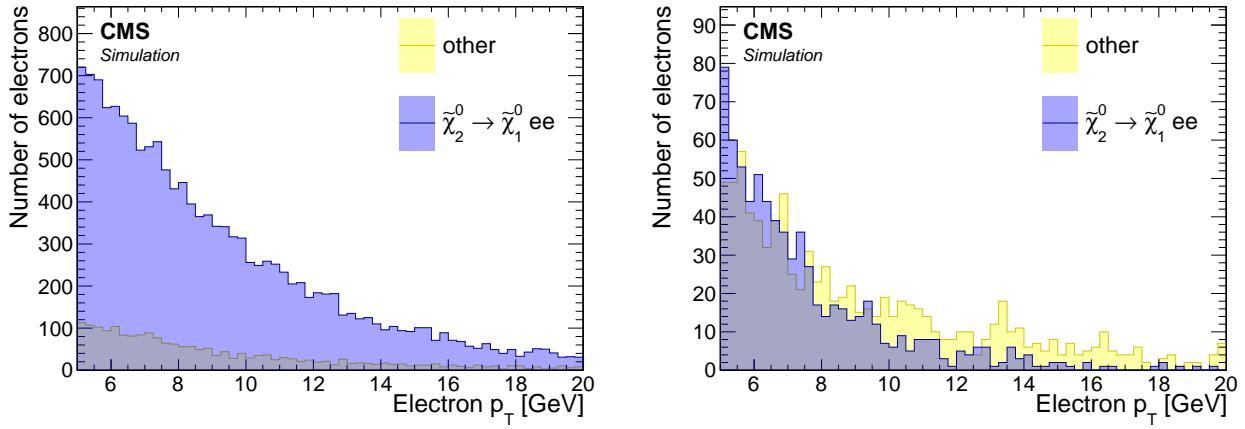


Figure 4.16: Distribution of reconstructed electron p_T with loose ID for $\Delta m = 5.63 \text{ GeV}$ (left) and $\Delta m = 1.92 \text{ GeV}$ (right). A cut of $\Delta R(j_1, e) > 0.4$ is applied.

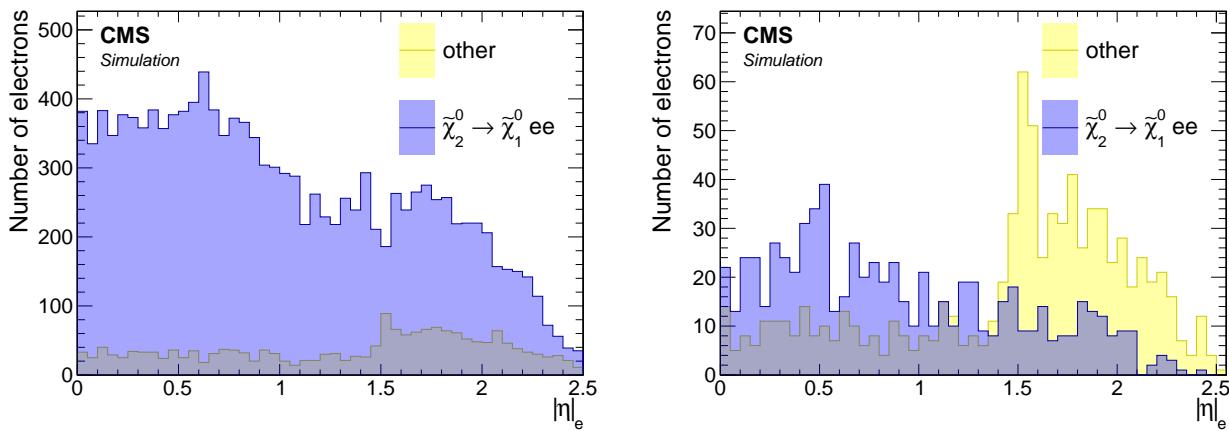


Figure 4.17: Distributions of $|\eta|$ of reconstructed electrons with loose ID for $\Delta m = 5.63 \text{ GeV}$ (left) and $\Delta m = 1.92 \text{ GeV}$ (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15 \text{ GeV}$ are applied.

To determine whether a tighter WP for the electron-identification is beneficial, the effects of requiring either a Medium or a Tight WP are investigated. The WP previously used in

the distributions is the loose WP. Two bins labeled *fail* and *pass* indicate the frequency with which the electron fails or passes the identification criteria of Medium or Tight WPs. These bins are shown in Figure 4.18. A considerable fraction of non-signal electrons are rejected in the low Δm case by picking either a Medium or Tight WP, but a significant number of signal electrons are also lost. Therefore, using these selections is not very efficient and results in low signal acceptance. The decision is made to use a loose WP for the electrons, and instead rely on isolation to achieve higher purity.

The effect of isolation on the purity of the electrons is also examined. The jet-based isolation is discussed in detail in Section 4.6.7, but for the sake of completeness, its effect on the purity of the electrons is also shown here. The custom jet-based isolation is compared with the standard definition of lepton isolation, which does not take into account the possibility that two electrons can be produced with a small angle of separation (small ΔR), as is the case for signal models with small Δm . The isolation distributions are shown in Figure 4.19.

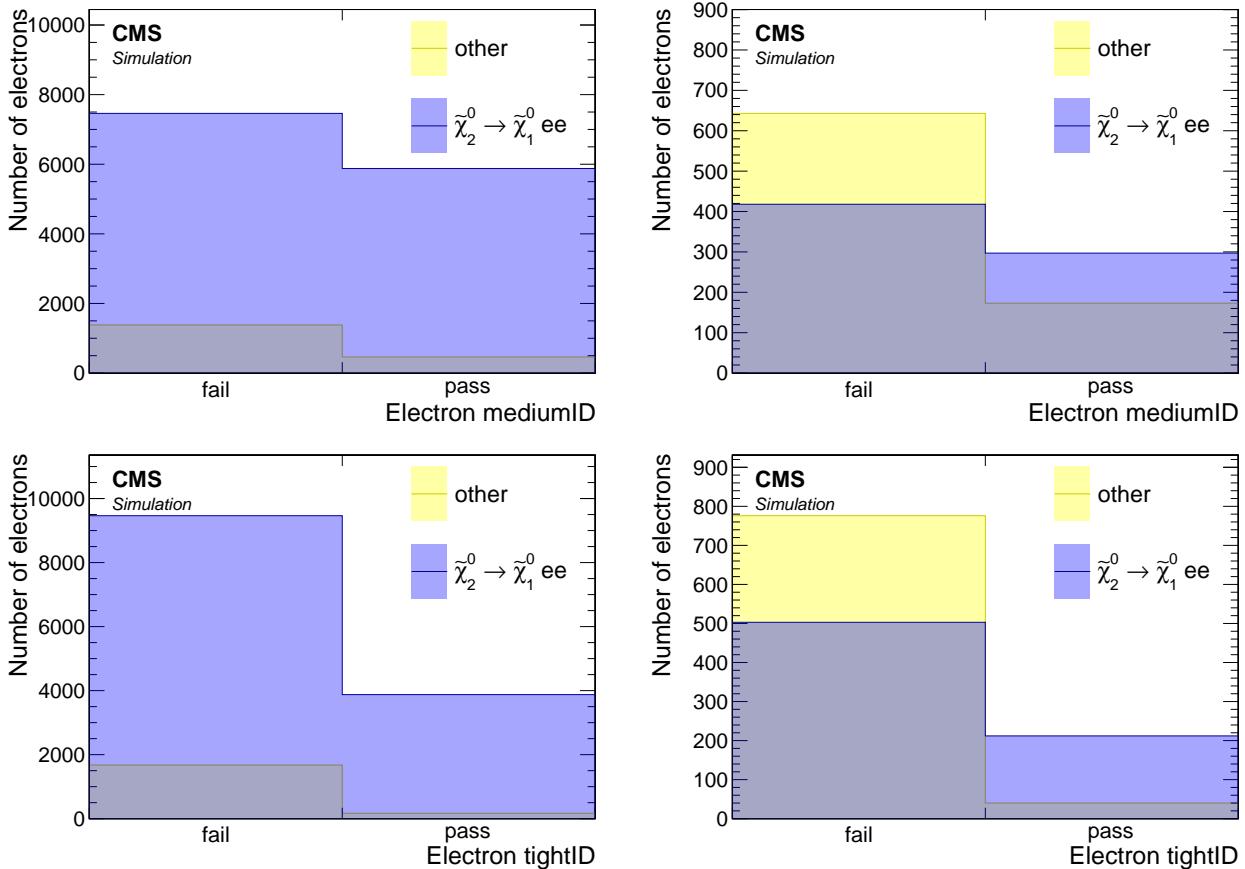


Figure 4.18: Medium (top) and Tight (bottom) ID WPs distributions of reconstructed electrons for $\Delta m = 5.63$ GeV (left) and $\Delta m = 1.92$ GeV (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15$ GeV are applied to the electrons.

The standard lepton isolation is not efficient for both Δm cases, while the custom jet-isolation performs well in terms of signal electron efficiency and successfully rejects a considerable amount of non-signal electrons. This results in a purer sample of electrons, and thus the choice of custom jet-isolation is concluded to be favorable. The effect of this choice on the η distribution is also examined in Figure 4.20, concluding the selection of electrons. The custom jet-isolation optimally purifies the electron sample while retaining a high signal efficiency, compared to distributions in Figure 4.17.

In summary, the following is the full set of selection criteria the analysis electrons:

- $5 < p_T < 15$ GeV;

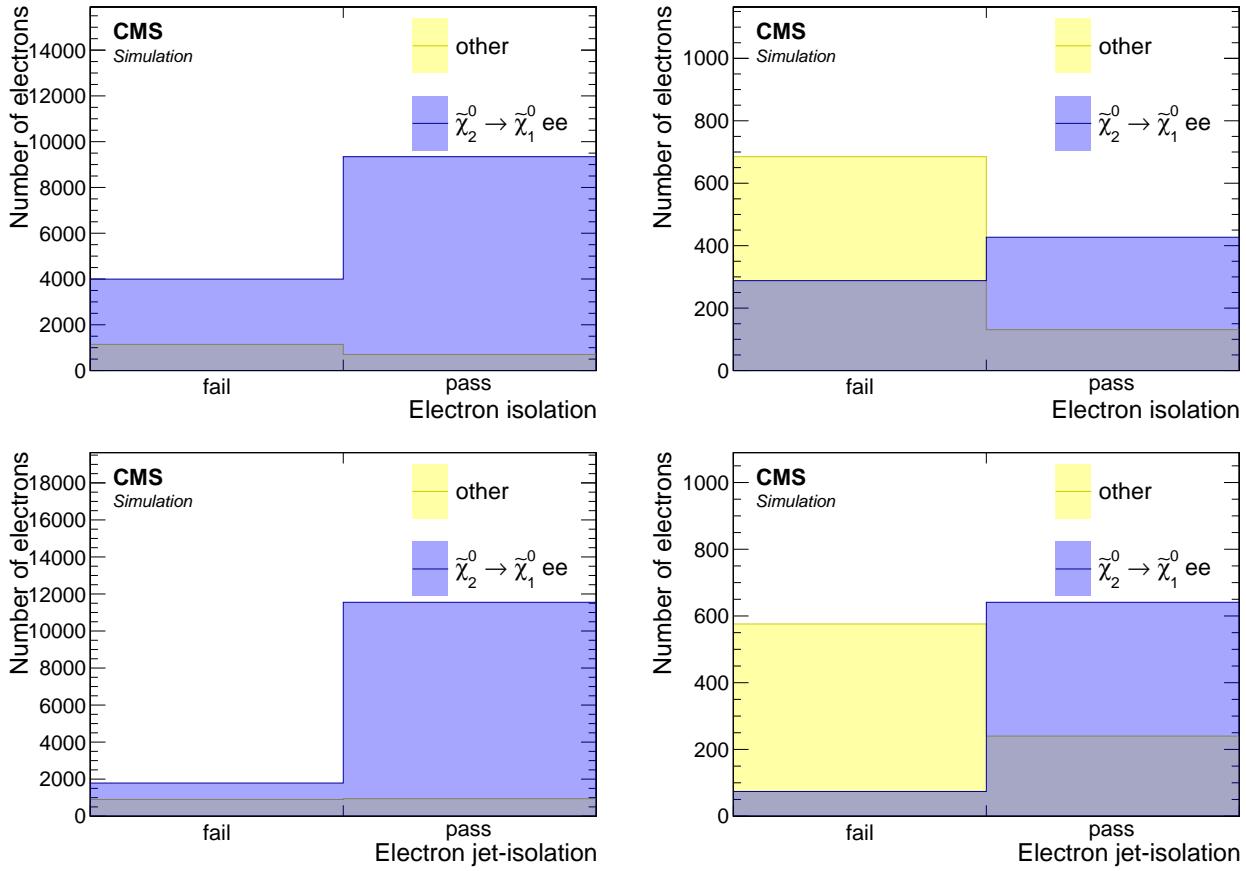


Figure 4.19: Standard isolation (top) and custom jet-isolation (bottom) distributions of reconstructed electrons with loose ID for $\Delta m = 5.63 \text{ GeV}$ (left) and $\Delta m = 1.92 \text{ GeV}$ (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15 \text{ GeV}$ are applied.

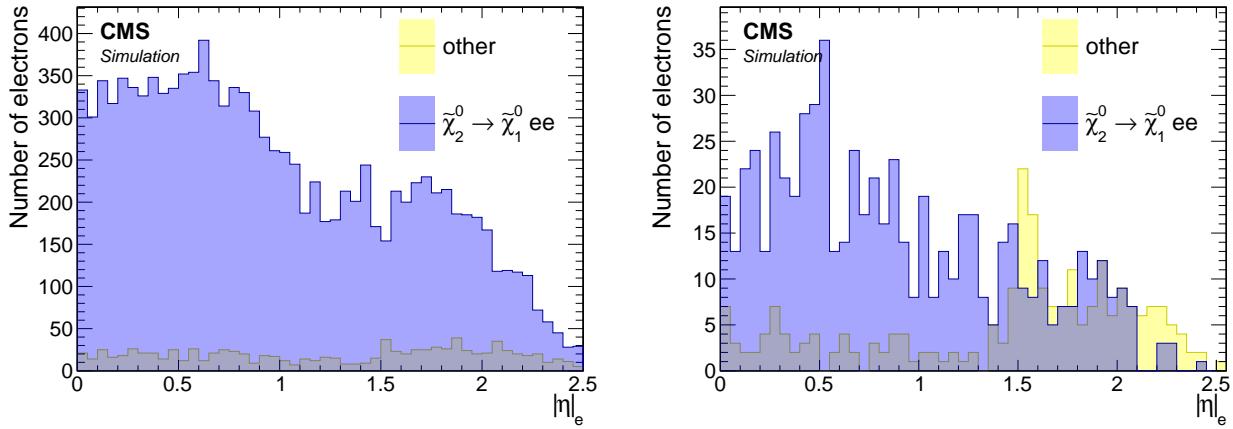


Figure 4.20: $|\eta|$ distribution of reconstructed electrons with loose ID passing jet-isolation for $\Delta m = 5.63 \text{ GeV}$ (left) and $\Delta m = 1.92 \text{ GeV}$ (right). Cuts of $\Delta R(j_1, e) > 0.4$ and $p_T < 15 \text{ GeV}$ are applied.

- $|\eta| < 2.5$;
- $\Delta R(j_1, e) > 0.4$;
- loose ID WP;
- pass jet-isolation.

4.6.2 Muons

The p_T threshold for reconstructed muons is significantly lower than that of electrons, making this channel particularly promising in terms of signal acceptance for low Δm models. As was the case for electrons, the initial WP choice for reconstructed muons is loose (more information in ??), and an analogous procedure is now followed for muons. The angular separation of muons from the leading jet in the event, $\Delta R(j_1, \mu)$, is the first distribution examined. As shown in Figure 4.7, the muon endcaps are capable of reconstructing muons with $p_T < 3$ GeV while the barrel is not. Therefore, a split view of barrel and endcaps is shown in Figure 4.21. Because the endcaps accept muons with lower p_T than the barrel, and because of the generally higher occupancy of tracks in the forward region, the purity in the endcaps is much lower than that in the barrel. The selection developed here attempts to further purify the somewhat contaminated barrel muon sample. Muons with $\Delta R(j_1, \mu) > 0.4$ are selected as in the electrons case, and this selection will apply for the rest of the section.

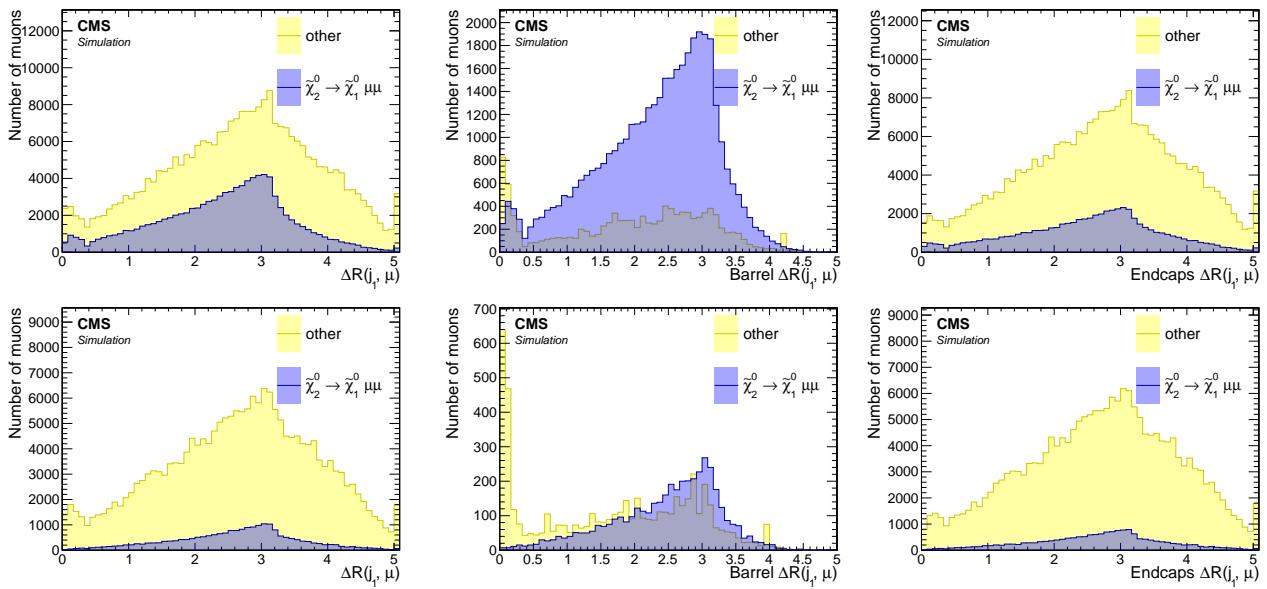


Figure 4.21: Angular separation between reconstructed muons with loose ID and the leading jet $\Delta R(j_1, \mu)$ for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive case (left), barrel (middle) and endcaps (right).

Distributions of muon p_T are examined having applied the previous cut of $\Delta R(j_1, \mu) > 0.4$. As seen in Section 4.4.4.1, the p_T distribution depends strongly on Δm . The p_T distributions seen in Figure 4.22 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 included as input to the multivariate classifier employed at a later stage, which can effectively cut tighter on the p_T dynamically and in concert with cutting on other variables. The actual maximum value of the p_T of the muons will depend on the Boosted Decision Tree (BDT) cut being used to define the signal region. p_T of the muons will depend on the BDT cut being used to define the signal region. The feature discussed earlier, whereby the endcaps are capable of reconstructing muons with lower p_T and therefore have worse purity than the barrel, is reiterated here. It is important to stress that the worse purity is due to a much higher efficiency, and as long as the muons can be purified further, it is not necessarily a bad thing. The rate of the non-signal muons in the region of $p_T < 2$ GeV, is seen to diverge rapidly, and the ratio of signal muons to non-signal muons is very low in that region. Therefore, an additional cut of $p_T > 2$ GeV is adopted. To evaluate the effect of this cut, the $|\eta|$ distribution before and after the p_T cut, is shown in Figure 4.23.

The impact of choosing an alternate WP, namely Medium or Tight, is examined in Fig-

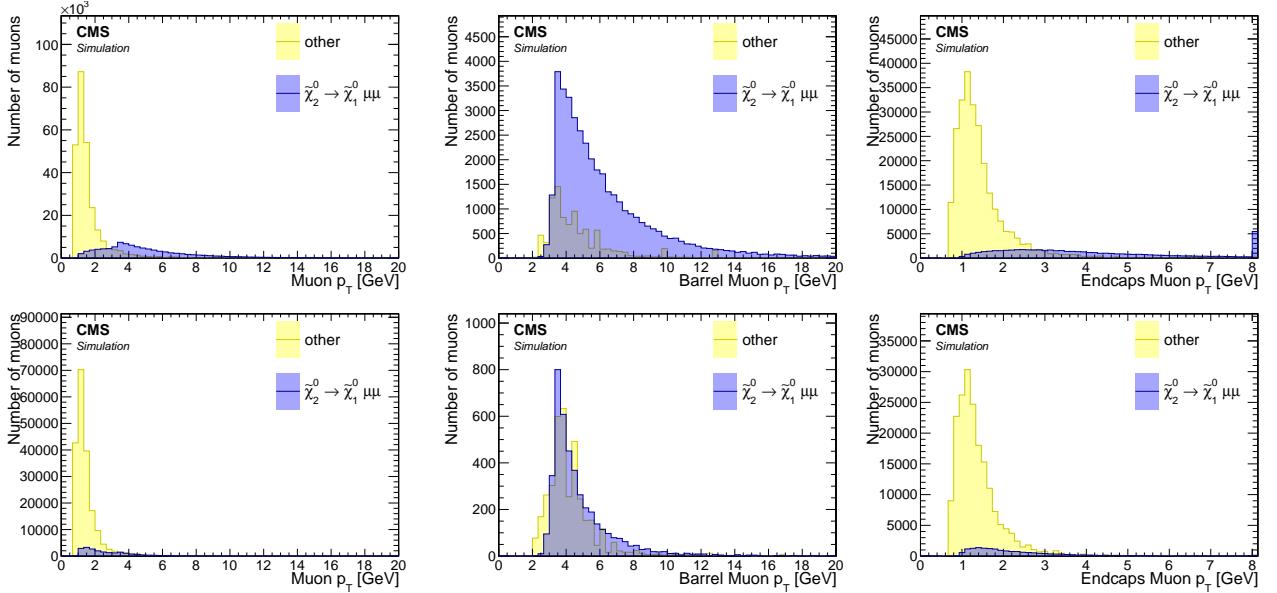


Figure 4.22: Distribution in signal events of the p_T of reconstructed muons with loose ID for $\Delta m = 5.63 \text{ GeV}$ (top) and $\Delta m = 1.92 \text{ GeV}$ (bottom) in the inclusive case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$ and $p_T < 15 \text{ GeV}$ are applied.

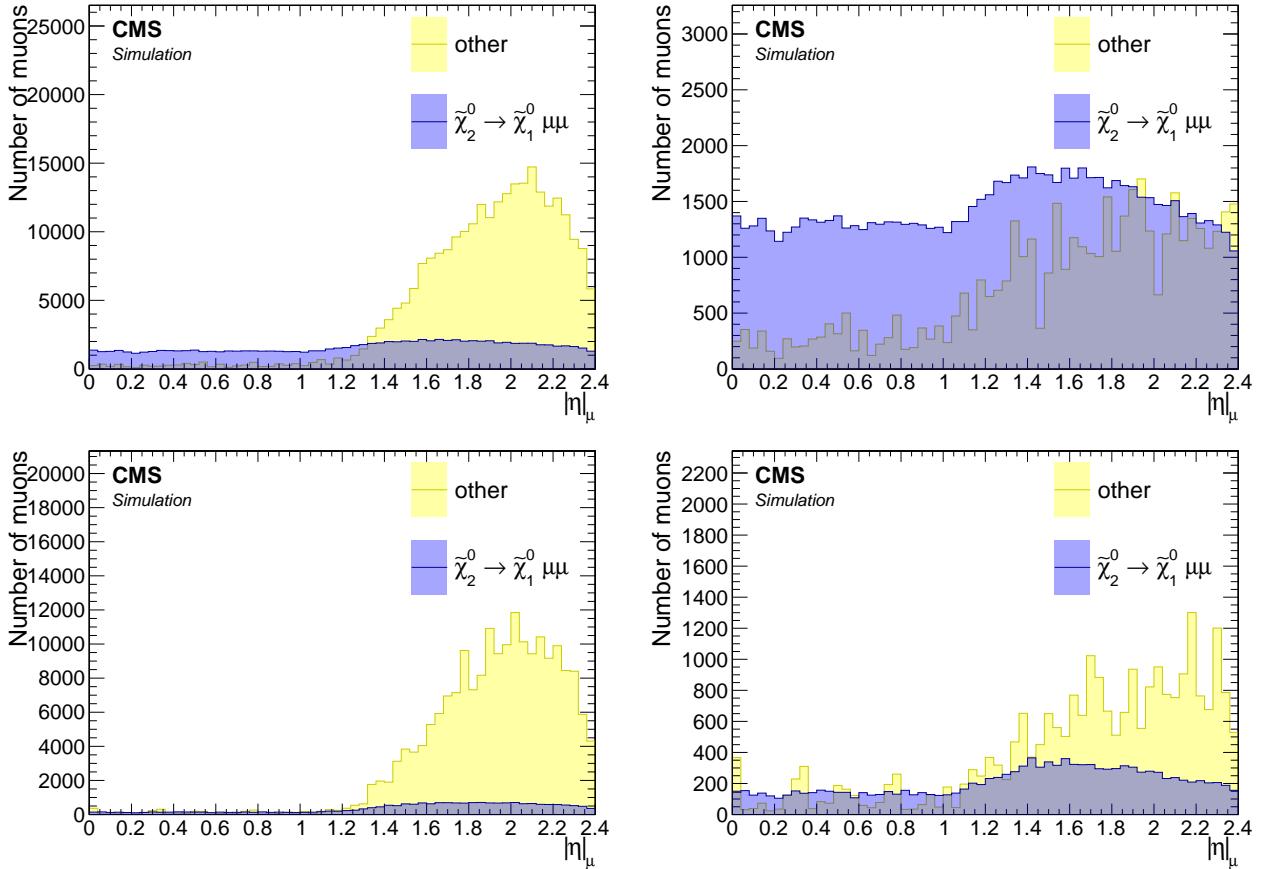


Figure 4.23: Distribution in signal events of the $|\eta|$ of reconstructed muons with loose ID for $\Delta m = 5.63 \text{ GeV}$ (top) and $\Delta m = 1.92 \text{ GeV}$ (bottom) without (left) and with (right) $p_T > 2 \text{ GeV}$ cut. A cut of $\Delta R(j_1, \mu) > 0.4$ is also applied.

ures 4.24 and 4.25, respectively. Two bins labeled *fail* and *pass* are plotted, which correspond to whether the muon passes or fails the identification criteria of a Medium or Tight WPs. The

Medium WP is seen to be highly performant in purifying the muon sample. The Tight WP on the other hand leads to a significant number of wanted signal-muons being lost without a significant gain in purity. Therefore, the Medium ID WP is chosen.

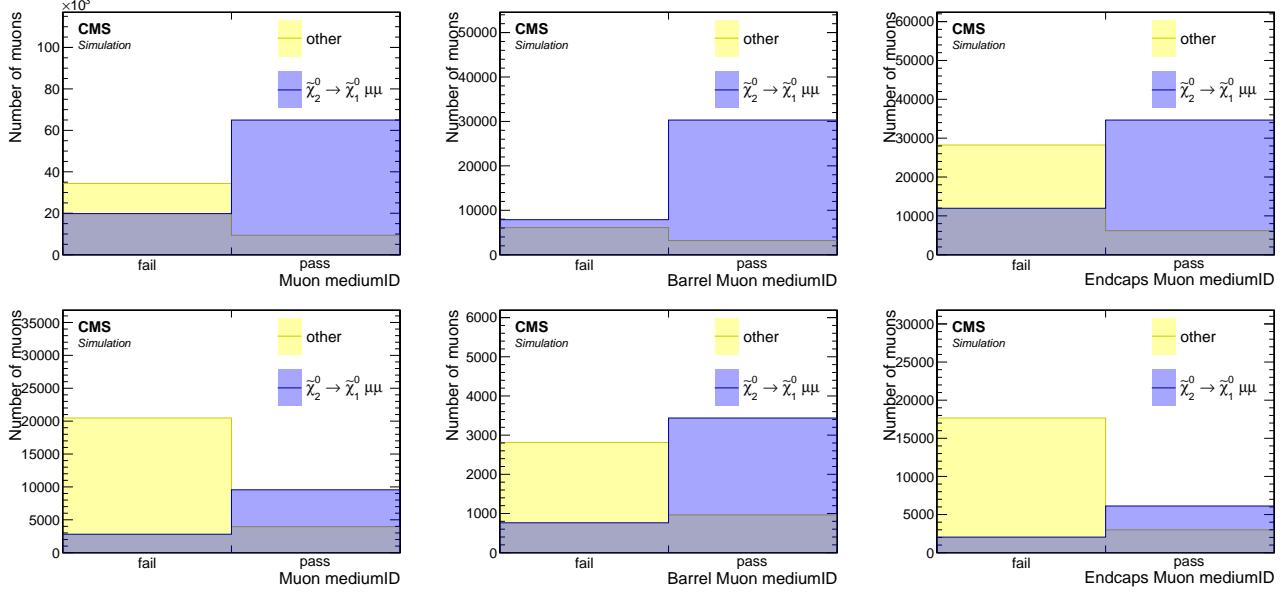


Figure 4.24: Medium ID WP distributions of reconstructed muons for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive p_T case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$, $p_T > 2$ GeV and $p_T < 15$ GeV are applied.

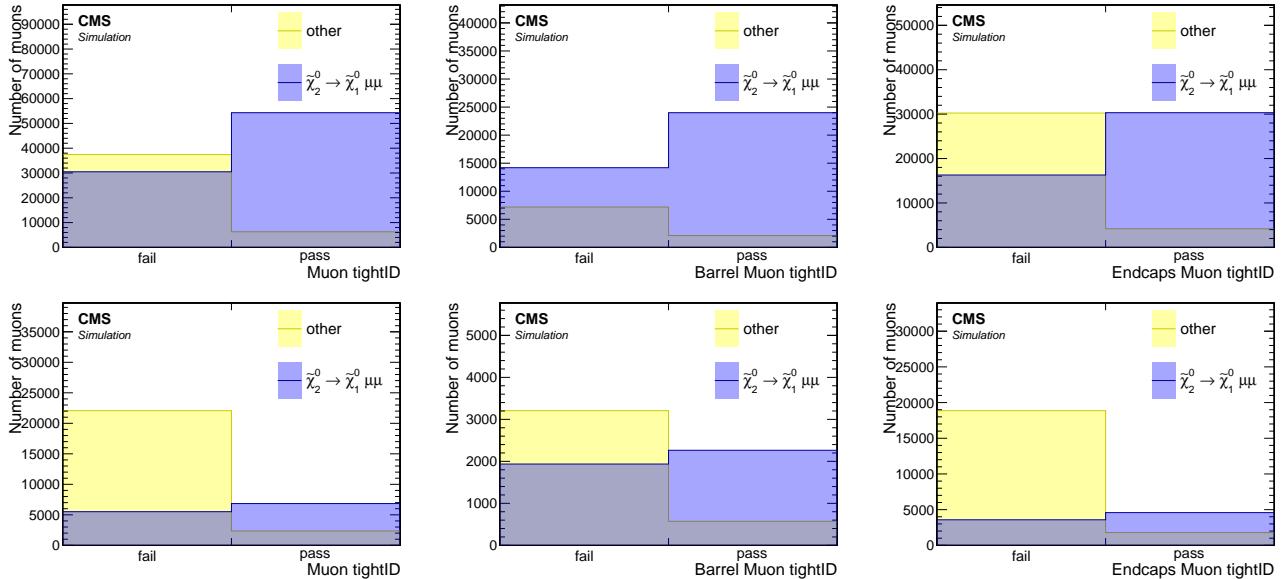


Figure 4.25: Tight ID WP distributions of reconstructed muons for $\Delta m = 5.63$ GeV (top) and $\Delta m = 1.92$ GeV (bottom) in the inclusive p_T case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(j_1, \mu) > 0.4$, $p_T > 2$ GeV and $p_T < 15$ GeV are applied.

The custom jet-isolation was designed to reject SM background while retaining signal, as the effects of the custom jet-based isolation, as described in Section 4.6.7, on signal muons is examined in this purity study. Figure 4.26 shows that a small price is paid by requiring the isolation. However, as will be seen in Section 4.6.7, the sensitivity is increased by rejecting a significant portion of SM background via the isolation criterion.

In summary, the following is the full set of selection criteria of the analysis muons:

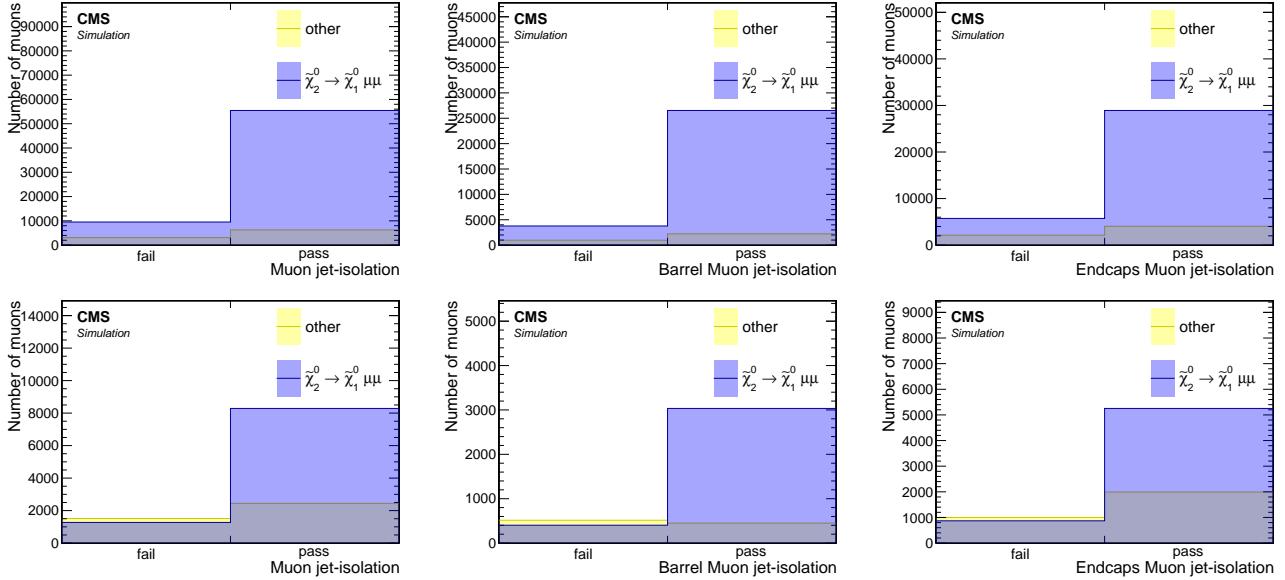


Figure 4.26: Distributions of the jet-based lepton isolation of reconstructed muons with Medium ID for $\Delta m = 5.63 \text{ GeV}$ (top) and $\Delta m = 1.92 \text{ GeV}$ (bottom) in the inclusive p_T case (left), barrel (middle) and endcaps (right). Cuts of $\Delta R(\gamma_1, \mu) > 0.4$, $p_T > 2 \text{ GeV}$ and $p_T < 15 \text{ GeV}$ are applied.

- $2 < p_T < 15 \text{ GeV}$;
- $|\eta| < 2.4$;
- $\Delta R(\gamma_1, \mu) > 0.4$;
- Medium ID WP;
- pass jet-isolation.

4.6.3 Scale factors

In Sections 4.6.1 and 4.6.2, the selection applied to electrons and muons was studied, and a choice was made regarding the identification working point used to draw conclusions about the identification efficiency of the leptons. However, relying solely on Monte Carlo (MC) can introduce systematic errors due to imperfections in modeling both the data and the detector response. Therefore, evaluating the level of consistency between data and MC is important. To this end, the lepton efficiency is studied in data and MC, where *efficiency* refers to the fraction of produced leptons that are reconstructed, selected and identified as such:

$$\varepsilon_\ell = \frac{N_\ell(\text{reconstructed, selected, and identified})}{N_\ell(\text{produced})} \quad (4.1)$$

In MC, the number of produced leptons is simply the number of generated leptons. However, in data, the efficiency must be measured in another way. Once the efficiency has been measured both in simulation and data, a correction factor called the Scale Factor (SF) can be applied as weights to the simulation to correct for discrepancies that may arise. Scale factors are defined as the ratio between the efficiency in data and the efficiency in simulation:

$$\text{SF}_\ell = \frac{\varepsilon_\ell^{\text{Data}}}{\varepsilon_\ell^{\text{MC}}} \quad (4.2)$$

The SFs are applied as weights for each lepton that passes the object selection in the event. The scale factors for loose-ID electrons in the relevant p_T range have been centrally measured by the corresponding working group and are applied to the selected electrons. Scale factors for needed (Medium) ID muons with $p_T \geq 2 \text{ GeV}$ were computed centrally by the Muon Physics Object Group (POG). However, as mentioned in [28, 29], the scale factors were computed by requiring $\Delta R > 0.5$, which excludes the key phase space that drives the sensitivity, as discussed in Section 4.4.4.3. Therefore, it is necessary to evaluate any potential ΔR dependence. To this end, the efficiency in different ΔR regions is extracted from an analysis of J/ψ decays.

To measure the lepton efficiency in data, tag and probe method is employed. This method often makes use of a known mass resonance such as Z , J/ψ or Υ to select particles using very loose selection criteria on one of the objects (the probe). The efficiency of applying one or more selection criteria is then computed as the fraction of probe particles that satisfy this criteria. The mass resonance results in the formation of a peaking structure in the invariant mass of the tag and probe system, from which the number of total and passing probe particles can be extracted via a sideband fit. To measure the efficiency of the analysis muons, muon/track pairs are selected with an invariant mass near the J/ψ mass. The muon is labeled as the ‘tag’ and the track as a ‘probe’. The tag muon is selected with a very tight selection resulting in high certainty that the object corresponds to a real produced muon. The probe (inner tracker track) corresponds to a very inclusive object, with no selection corresponding to the muon ID applied to it, a loose selection but constrained to be consistent with a product of a J/ψ . The background contribution, originating from random tracks as well as fake muons, is removed by simultaneous fit of a smoothly falling background and the shape of the J/ψ . The probes are then subjected to the requirement that the track be matched within $\Delta R < 0.01$ of an analysis muon. The efficiency in question is written as:

$$\varepsilon_\mu = \frac{N_\mu^{\text{Matched}}}{N_t} \quad (4.3)$$

The selection applied on the objects corresponding to the numerator and denominator are summarized in Table 4.6.3. This study was conducted for the year 2016. The 2016 samples listed in Section 4.5.1 are used for MC. To ensure the independence of the tagged muon from

the triggered object, a single electron trigger is used for data. The corresponding data set is measured to be 36.02 fb^{-1} using the BRIL Work Suite [30]. The following trigger paths are utilized:

- HLT_Ele27_WPTight_Gsf_v*,
- HLT_Ele27_eta2p1_WP Loose_Gsf_v*,
- HLT_Ele32_WPTight_Gsf_v*,
- HLT_Ele35_WPTight_Gsf_v*.

An offline loose ID electron with $p_T > 27 \text{ GeV}$ is then selected. The requirements to select a tag and probe pair are defined in Table 4.6.3.

Table 4.3: 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 performed in an invariant mass window around the J/ψ peak of $[2.5, 3.5] \text{ GeV}$. The signal PDF is modelled using a double-sided Crystal Ball function, and the continuum is modelled with a 6th order polynomial. The fit is repeated twice, where the denominator is based on probe tracks, and the numerator uses those tracks that have been matched to Medium ID muons. The ΔR range has been split into three, and the $|\eta|$ of the muons has been divided into barrel ($|\eta| < 1.2$) and endcaps ($1.2 < |\eta| < 2.4$). Fits to the MC are shown in Figure 4.27 for the barrel and Figure 4.28 for the endcaps. Fits to the data are shown in Figure 4.29 for the barrel and Figure 4.30 for the endcaps.

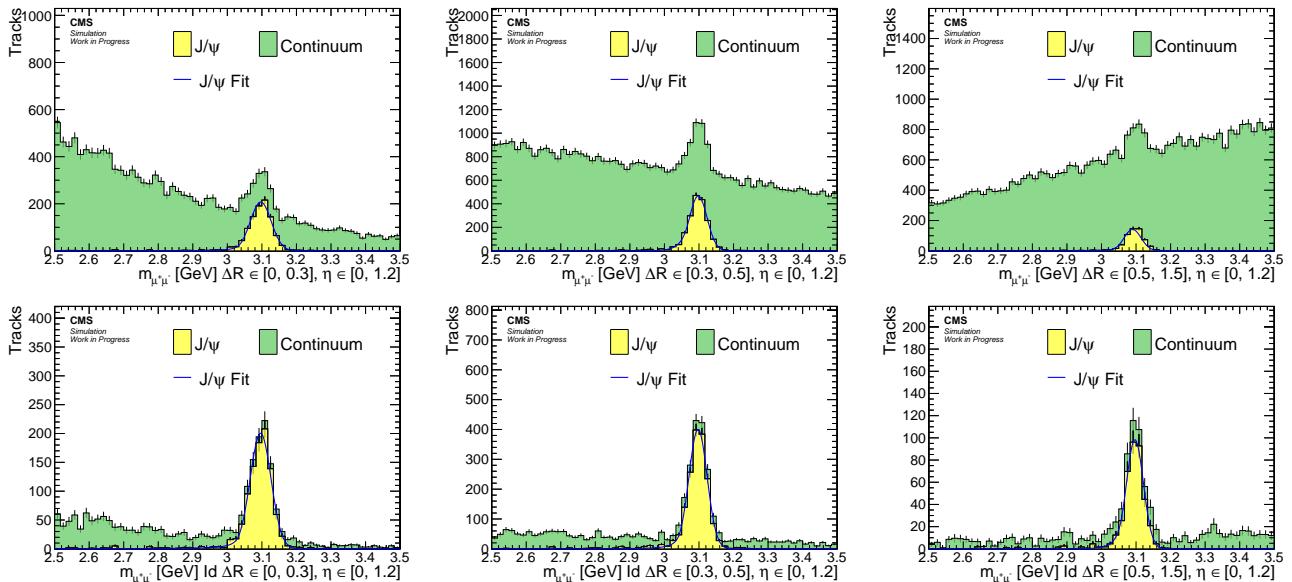


Figure 4.27: Fits to the tag and probe invariant mass for muons in the barrel region based on MC. Results are shown for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right).

The efficiency and corresponding scale factors are shown in Figure 4.31. The scale factors are statistically consistent with unity and show no discernible ΔR dependence. A similar study

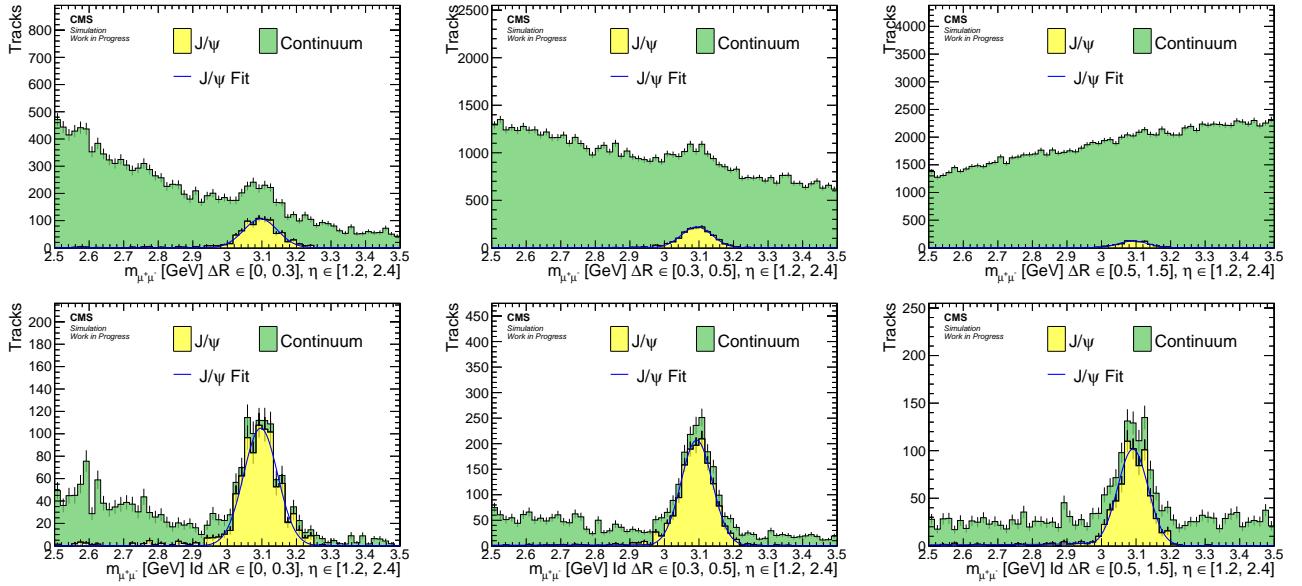


Figure 4.28: Fits to the tag and probe invariant mass for muons in the endcaps region based on MC. Results are shown for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right).

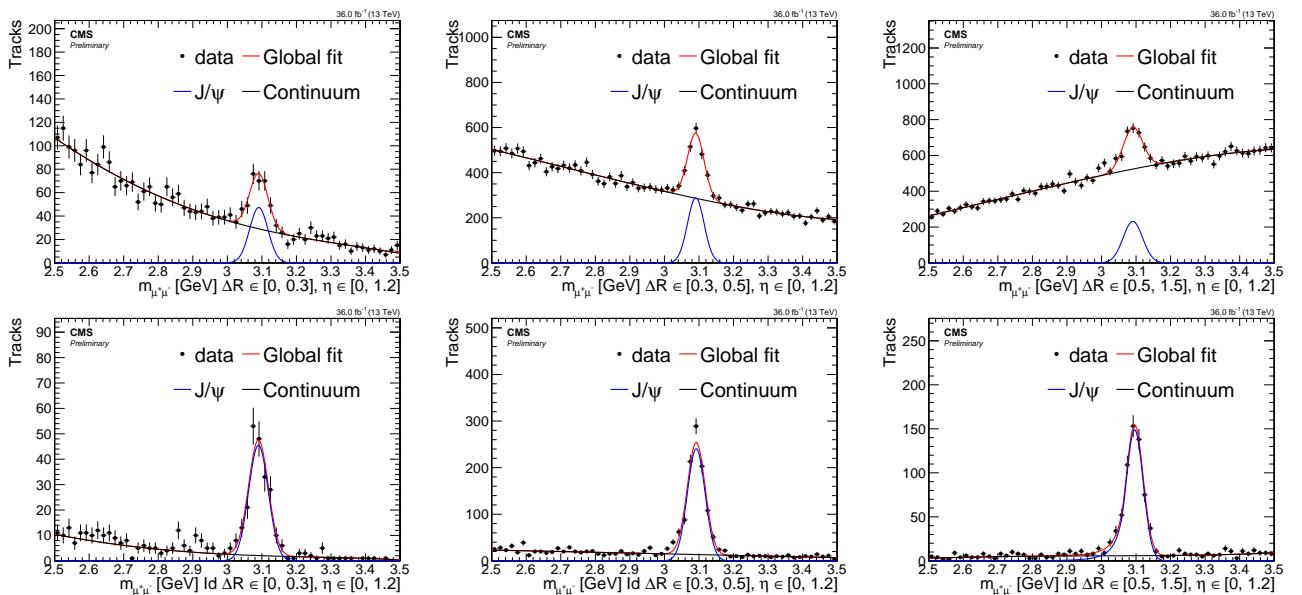


Figure 4.29: Fits to the tag and probe invariant mass for muons in the barrel region based on data. Results are shown for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right).

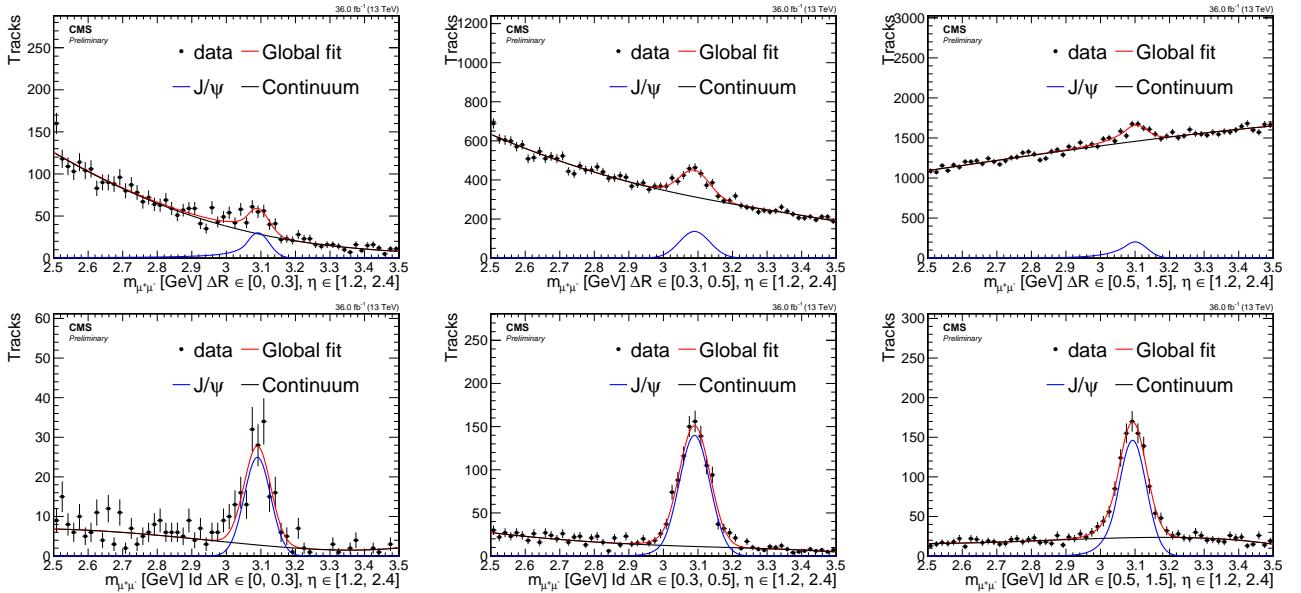


Figure 4.30: Fits to the tag and probe invariant mass for muons in the endcaps region based on data. Results are shown for denominator (top) and numerator (bottom) for $0 < \Delta R < 0.3$ (left), $0.3 < \Delta R < 0.5$ (center), $0.5 < \Delta R < 1.5$ (right).

was carried out based on 2017 and 2018 data and MC [31], and no ΔR dependence was observed either. As a result of these studies, the recommendation from the POG is to use the calculated scale factors provided by them with an additional systematic uncertainty of 1% for muons with $p_T < 20 \text{ GeV}$.

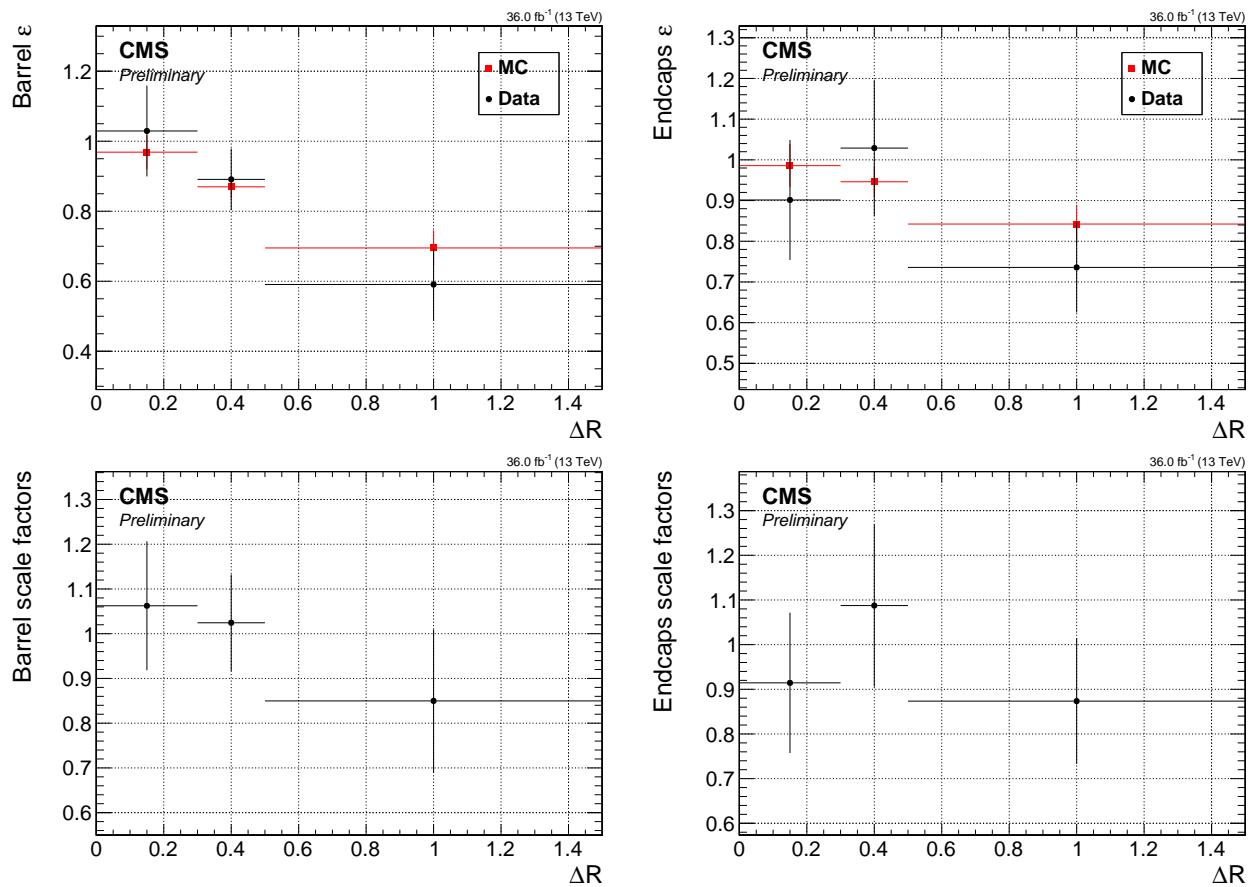


Figure 4.31: Efficiencies (top) and scale factors (bottom) for barrel muons (left) and endcaps muons (right).

4.6.4 Missing transverse energy

The importance of the missing transverse momentum (or energy) in this analysis has been discussed in Section 4.4.1. Two standard measures of the momentum imbalance in the events are \vec{E}_T^{miss} (or \vec{p}_T^{miss} by a different symbol) and \vec{H}_T^{miss} . \vec{E}_T^{miss} is defined as:

$$\vec{E}_T^{\text{miss}} = \vec{p}_T^{\text{miss}} = - \sum_i \vec{p}_T(i), \quad (4.4)$$

where the summation is done on all particle flow candidates. Therefore, the missing transverse energy serves as a measure of particles that evade detection, such as weakly interacting neutral particles. Mismeasurements of visible particles and additional energy deposits from sources such as Pile-Up (PU), jet energy response and detector noise can affect this observable, so it undergoes further correction to mitigate these effects. The correction process considers jets with p_T greater than 10 GeV. Full details of the corrections can be found in [32].

An alternative measurement to the missing transverse momentum is \vec{H}_T^{miss} , which is sometimes referred to as *missing hardronic activity*. Instead of considering all particle flow candidates in the sum, this measurement only takes into account jets with p_T greater than 30 GeV and $|\eta|$ less than 5, and is defined as:

$$\vec{H}_T^{\text{miss}} = - \sum_i^{\text{jets}} \vec{p}_T(i). \quad (4.5)$$

The observable \vec{H}_T^{miss} is favored over \vec{E}_T^{miss} in this analysis because the jet-based isolation, defined in Section 4.6.7, uses jets with p_T greater than 30 GeV, while a sideband is defined using the range of $p_T \in [15, 30]$ GeV of jets, which is then used for the estimation of the jetty background in Section 4.9.2.1. Both observables, \vec{E}_T^{miss} and \vec{H}_T^{miss} , have equivalent scalar quantities, E_T^{miss} and H_T^{miss} respectively, which can be obtained by taking the magnitude of their vectorial counterpart.

4.6.5 Jets

Jets used in the analysis are reconstructed by clustering the Particle Flow (PF) candidates using FASTJET with the anti- k_T algorithm [33] with a size parameter of 0.4. Tagging of b -jets is performed using the multivariate technique DEEPCSV with a Medium WP, also known as the Combined Secondary Vertex (CSV) algorithm [34]. Jets are required to have a transverse momentum $p_T > 30$ GeV and $|\eta| < 2.4$.

4.6.6 Tracks and multivariate selection

The leptons $\ell^+ \ell^-$ produced in the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$ tend to have very low transverse momentum p_T . It was shown in Section ?? that the identification and reconstruction efficiency of the muons worsens with lower p_T . Therefore, the aim of the exclusive track category is to recover lost leptons that were not reconstructed or identified. As seen in Section ??, the tracking efficiency for the p_T ranges used in this analysis is well above 99%, allowing the recovery of some of the tracks that correspond to the missing leptons.

To identify which track corresponds to the target lepton in a given signal event, a BDT classifier is trained. Four separate BDTs are trained, corresponding to each lepton flavor (muon or electron) and each phase of the tracker (Phase 0 for 2016, and Phase 1 for 2017-2018). All BDTs use a common structure of 200 trees with a maximum depth of 3, and are trained with AdaBoost and GiniIndex separation using the TMVA package [35]. The package's default values are used for all other parameters. Tracks from a dedicated FASTSIM signal simulations described in Section 4.5.2 are used for training. A broad range of simulated higgsino parameter μ (or the mass of $\tilde{\chi}_1^\pm$) is considered, but only the range of Δm that this analysis targets. For Phase 0, Δm^0 is chosen from the range $[0.3, 4.3]$ GeV and μ from $[100 - 130]$ GeV, while for Phase 1, Δm^\pm is chosen from $[0.3 - 4.6]$ GeV and μ from $[100 - 500]$ GeV. Signal events are split into signal tracks and background tracks, with signal tracks originating from leptons from the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^-$ while background tracks do not match to the leptons. 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 independent samples of equal size. Distributions of the classifier score for signal and background categories with the test samples overlayed onto the training samples can be seen in Figure 4.33.

Pre-selection is applied to all tracks in the collection obtained by the standard track reconstruction sequences. This pre-selection ensures that only properly-reconstructed, isolated, and prompt tracks are considered. The selected tracks must also have trajectories passing through the region near the primary vertex (PV) with the largest sum of charged-tracks, jets, and missing energy values. The full set of track pre-selection criteria are

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

For the training, a set of 10 variables, listed in decreasing order of their importance ranking in Table 4.6.6 (in the muon case of Phase 0) is used.

Figure 4.32 shows the distribution of input variables, where signal tracks are shown in blue and background tracks in red.

The classifier output score for the 4 BDTs is displayed in Figure 4.33, where the test distributions are superimposed on the training sample. No obvious over-training is observed. The ROC curves are plotted in Figure 4.34, where the red point indicates the efficiency of the signal and background tracks of the minimum BDT cut, which is taken to be 0.0. Good separation between signal tracks and fake tracks is obtained, as evidenced by the relatively high

Table 4.4: Input variables to the in-signal track selecting classifier.

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}}) $	
5	$ \Delta\eta(t, j_1) $	j_1 is the leading jet
6	$ \Delta\phi(t, \ell) $	
7	$ \eta(t) $	
8	$ \eta(\ell) $	
9	$\Delta R(\ell, j_1)$	
10	$m_{t\ell}$	invariant mass

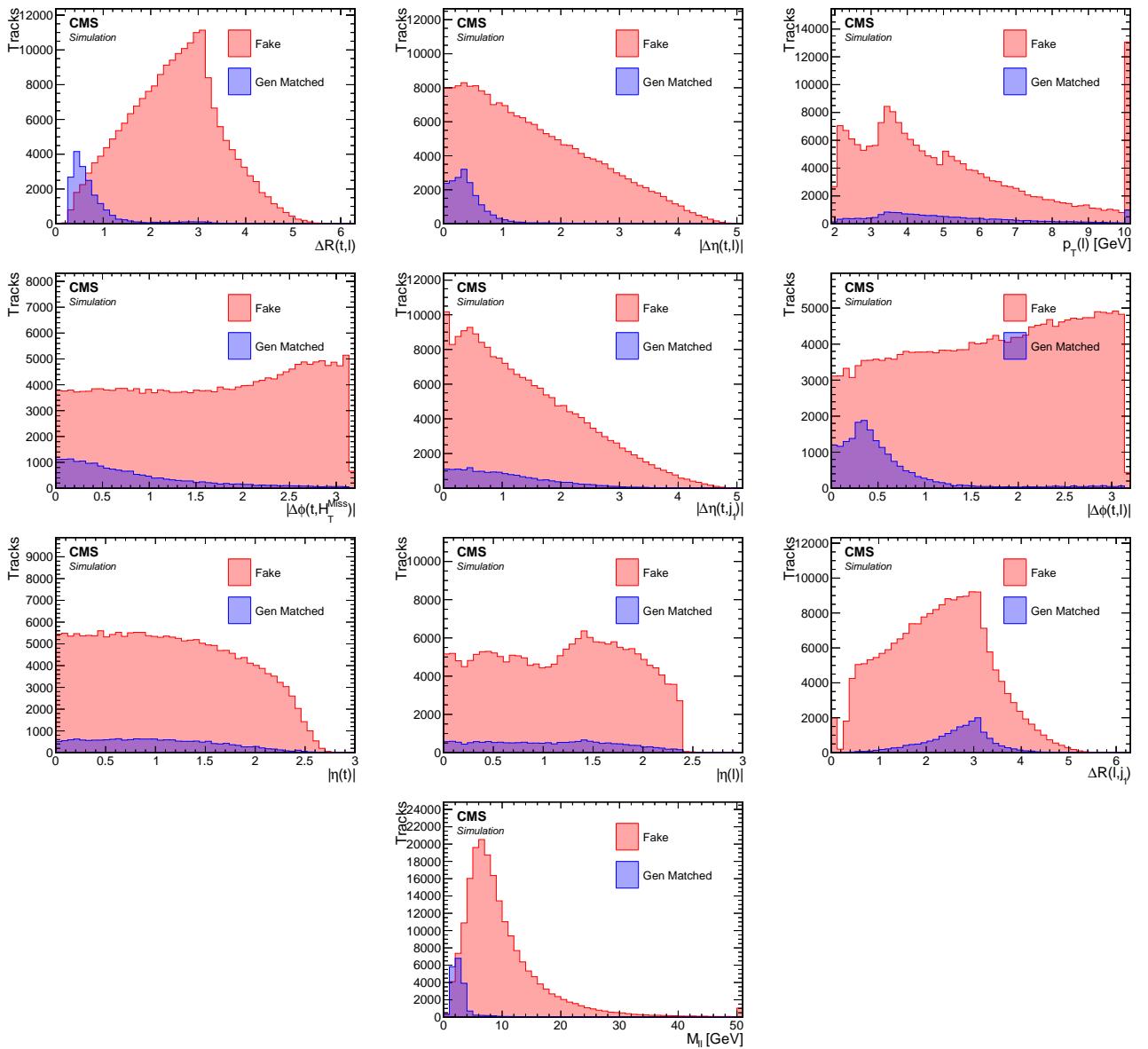


Figure 4.32: Distributions of the inputs to the track BDT in the muon exclusive track category. Fake category refers to tracks not originating from target leptons.

signal efficiency of over 90% (86%) for muons (electrons) and background rejection of around 86% (76%) for muons (electrons).

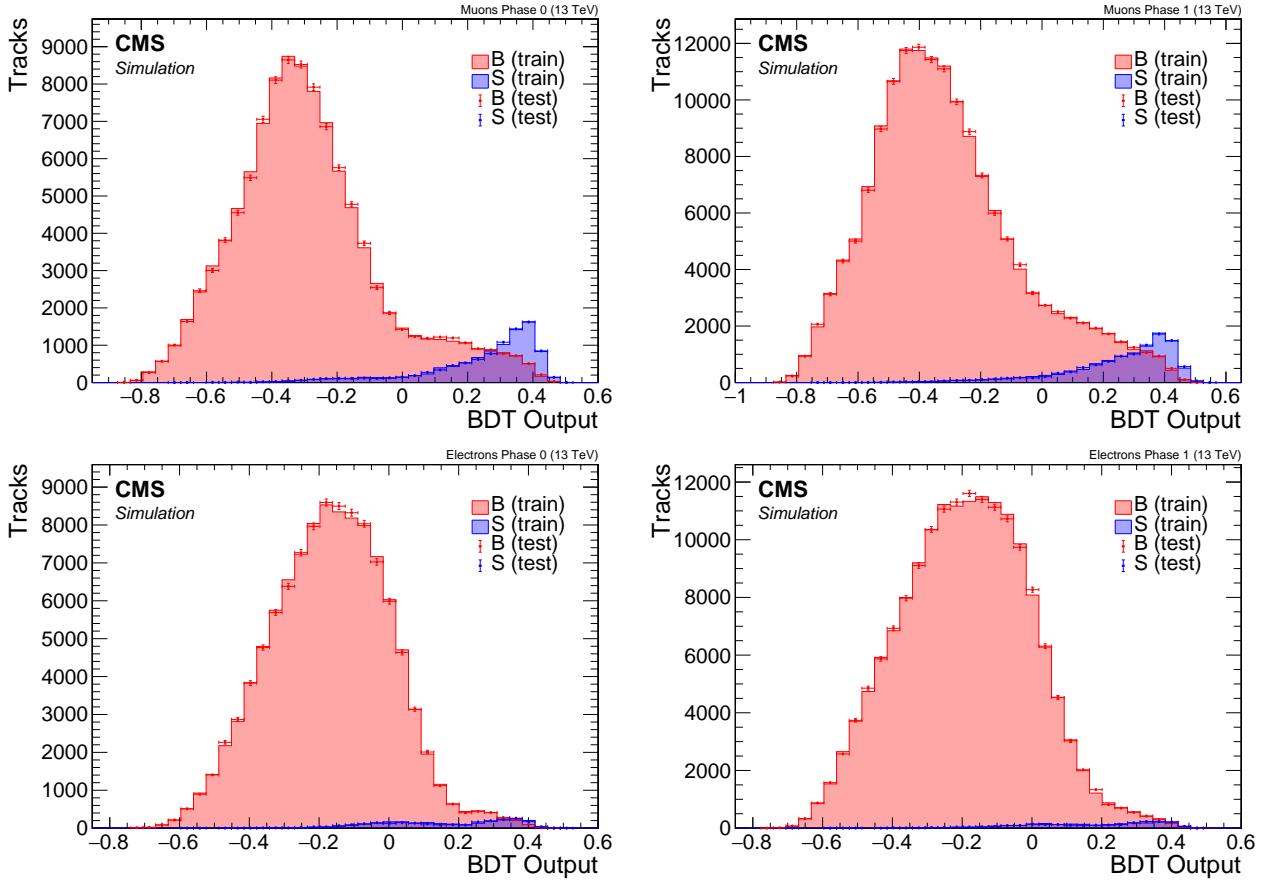


Figure 4.33: Track BDT output plots for Muons (top) and Electrons (bottom) in Phase 0 (left) and Phase 1 (right). Blue shows signal tracks, while Red are fake tracks. Test sample overlay on top of training sample.

The track with the maximum BDT score is selected as the signal candidate track. Only events with a track with a score greater than 0.0, corresponding to the red dot in the ROC curves shown in Figure 4.34, are considered.

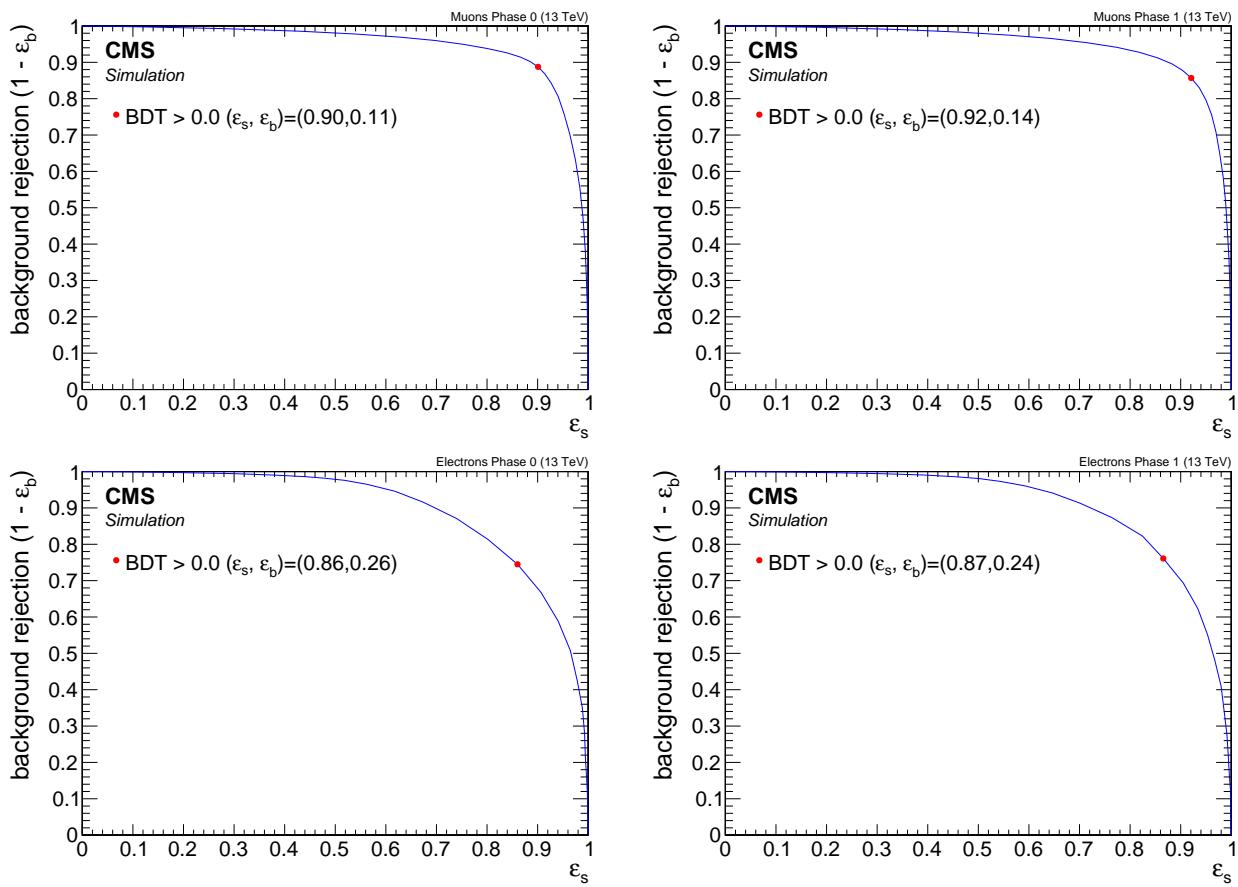


Figure 4.34: Track BDT ROC curves for muons (top) and electrons (bottom) in Phase 0 (left) and Phase 1 (right). The minimum threshold on the classifier score is indicated by the red dot.

4.6.7 Isolation

The leptons 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 in their vicinity. This is because the only jets in the event come from initial state radiation, which boosts the produced electroweakinos in the opposite direction. Therefore, the leptons originating from those electroweakinos will not propagate collinear to these jets. The characteristic signal event topology can be exploited to distinguish signal events from background originating from SM processes. At CMS, various standard isolation criteria are used. The three most widely used isolation criteria are track relative isolation [36], PF relative isolation (RelIso), which was first described in [37], and a modified version referred to as relative mini-isolation (miniRelIso), described in [38].

Track relative isolation is defined as the p_T sum of all tracks around a given track (or lepton) within an angular separation ΔR of 0.3:

$$\text{Track relative isolation}_{\ell} = \frac{\sum_{\substack{\text{tracks from PV} \\ \text{in } \Delta R < 0.3}} p_T}{p_T(\ell)}. \quad (4.6)$$

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

$$\text{RelIso}_{\ell} = \frac{\sum_{\substack{\text{charged} \\ \text{hadrons} \\ \text{from PV}}} p_T + \max \left(0, \sum_{\substack{\text{neutral} \\ \text{hadrons}}} E_T + \sum_{\text{photons}} E_T - 0.5 \cdot \sum_{\substack{\text{charged} \\ \text{hadrons} \\ \text{from PU}}} p_T \right)}{p_T(\ell)}. \quad (4.7)$$

The last term in the definition is a correction for PU effects. A lepton is considered to be isolated if its RelIso value is small. A variant of the relative isolation is the so-called mini relative isolation (miniRelIso), which differs from the standard relative isolation in that its cone size is dependent on the p_T of the lepton, as follows:

$$\Delta R = \begin{cases} 0.2 & p_T(\ell) \leq 50 \text{ GeV} \\ \frac{10 \text{ GeV}}{p_T(\ell)} & p_T(\ell) \in (50 \text{ GeV}, 200 \text{ GeV}) \\ 0.05 & p_T(\ell) \geq 200 \text{ GeV} \end{cases}. \quad (4.8)$$

The variable size cone allows for the recovery of efficiency when leptons are produced in the decay chain of a boosted object. In such cases, when the boost is large, the lepton is likely to overlap with another lepton produced at a common decay vertex, failing a standard isolation cut. The parameters are tuned to, and thus well-suited to, leptons from the decay of on-shell W and Z bosons, but are not suitable for low-mass resonances.

The drawback of standard isolation criteria in the case of this analysis's signal, the leptons can compromise each other's isolation. As shown in Section 4.4.4.3, access to low Δm model-points requires including the $\Delta R < 0.3$ phasespace region. Requiring any of the standard isolation criteria will thus result in rejecting valuable signal events. An alternative isolation criterion is proposed to help retain some of the desired phasespace while rejecting the majority of the standard model background. This alternative isolation proves to be useful not only for optimally selecting leptons, but also for defining a sideband control region needed for the jetty background estimation, as described in Section 4.9.2.1. The steps to construct the alternative *jet-based isolation* are described algorithmically below.

1. Subtract the vector 4-momenta of candidate leptons of a given flavor from any reconstructed jet ΔR smaller than 0.4

2. The lepton is said to pass isolation if it does not lie within $\Delta R < r$ of any *lepton-corrected* jet with $p_T > p$.
3. Lepton is said to fail isolation for background estimation if it fails *jet-based isolation*, and the nearest uncorrected jet has $15 < p_T < 30 \text{ GeV}$ (see 4.9.2.1 for use of such lepton)

The main idea behind defining jet-isolation is to reject leptons with hadronic activity around them while not losing a lepton that is close to another lepton of the same flavor. The process described introduces two free parameters: the p_T threshold of the lepton-corrected jets that cause a lepton's isolation to fail (p), and the cone size (r), which determines how close a corrected jet is allowed to be to a lepton. To choose the thresholds for these parameters, a scan is performed ranging over $p \in [0, 20] \text{ GeV}$ and $r \in [0.4, 0.6]$. For each step in the scan, the full analysis is performed, including the background estimation procedure which makes use of the jet-based isolation, and various performance criteria are extracted to inform the choice of optimum r and p . The criteria of interest include signal efficiency (which should be high), background efficiency (which should be low), signal contamination in 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 should be maximized). The scan is carried out for muons using 2016 MC and data, and the results are shown in Tables 4.6.7- 4.6.7.

Table 4.5: Signal efficiency for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.

	r				
	0.4	0.45	0.5	0.55	0.6
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

From Table 4.6.7, it is evident that the transfer factor of the jetty background estimation method increases with larger p and with smaller r . A transfer factor that is less than unity is preferred in order to ensure a high likelihood of well-populated control regions, and choices that do not meet this criterion are excluded. After taking into account all factors, the values $(p, r) = (10 \text{ GeV}, 0.6)$ are selected for muons and $(p, r) = (10 \text{ GeV}, 0.5)$ for electrons.

Table 4.6: Background efficiency for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.

	r				
	0.4	0.45	0.5	0.55	0.6
p	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	0.15
	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
					0.37

Table 4.7: Transfer factor for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.

	r				
	0.4	0.45	0.5	0.55	0.6
p	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
					3.86

Table 4.8: Significance $s/\sqrt{b + \epsilon_b^2}$ for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.

		r				
		0.4	0.45	0.5	0.55	0.6
p	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	0.23
	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

4.7 Trigger

4.8 Event selection

As discussed in Section 4.3, three event categories are used in this analysis: the dimuon category and an exclusive track category for each lepton flavor (muon and electron). The preselection is summarized in Section 4.8.1, followed by the selection that defines each category in Section 4.8.2. Finally, the multivariate selection for each category is discussed in Section 4.8.3.

4.8.1 Preselection

In Section 4.4.3, the preselection criteria that apply to all categories was defined. This section reiterates the reasons for this selection as well as describes other event-level selection.

- $H_T^{\text{miss}} \geq 220 \text{ GeV}$ and $E_T^{\text{miss}} \geq 140 \text{ GeV}$ cuts are intended to boost sensitivity by rejecting SM background and to operate in the acceptance regime of the MET trigger, as described in Section 4.7. These cuts are especially efficient in rejecting QCD background, which does not produce real E_T^{miss} . Any E_T^{miss} apparent in QCD is due to jet energy miss-measurements. The harder cut on H_T^{miss} is made instead of E_T^{miss} because H_T^{miss} sums jets with $p_T > 30 \text{ GeV}$ and is blind to objects with $p_T < 30 \text{ GeV}$. Background estimation relies on jets with p_T in the range of $[15, 30] \text{ GeV}$, so H_T^{miss} avoids introducing bias in the data-driven background estimation methods.
- $N_{\text{jets}} (p_T \geq 30 \text{ GeV} \text{ and } |\eta| < 2.4) \geq 1$. At least one jet is required in the event because such an ISR jet gives a boost to the produced neutralino, thus increasing the missing transverse energy and the sensitivity of the analysis.
- $N_{\text{b-jets}} (p_T \geq 30 \text{ GeV} \text{ and } |\eta| < 2.4) = 0$. Any event with b-tagged jet is vetoed since 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.
- $\min \Delta\phi (\vec{H}_T^{\text{miss}}, \vec{j}) > 0.4$. Requiring an ISR jet in the event leads to the expectation that the H_T^{miss} should point in the opposite direction of the jet or at an angle close to π . Events with multiple jets in the SM background, such as those arising from QCD, will not exhibit such a feature. Therefore, this cut reduces the QCD background.
- veto events with isolated loose-ID lepton having $p_T \geq 30 \text{ GeV}$. Lepton can be either muon or electron.
- $0.4 < m_{\ell\ell} < 12 \text{ GeV}$. The 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 be further tightened by the boosted decision tree.

The object level selection was described in detail in Section 4.6. For the sake of completeness it is reiterated. The electrons in the analysis require are required to pass the following selection (also described in Section 4.6.1):

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

The muons in the analysis are required to pass the following selection (see also Section 4.6.2):

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

4.8.2 Category selection

The analysis includes three main categories: the dilepton category and an exclusive track category for each lepton flavor. The dilepton category requires two fully-identified leptons, both of which are muons. In contrast, the exclusive track category includes a single lepton and a track that has not been identified as a lepton. Both electrons and muons are accepted as the single lepton in the exclusive track category. The selection criteria for the dilepton category are described in Section 4.8.2.1, while those for the exclusive track category are detailed in Section 4.8.2.2.

4.8.2.1 Dilepton selection

In the dilepton category, two reconstructed and identified muons are required. Events in the dilepton category must satisfy the preselection and the baseline selection, as well as the following criteria:

- $N_\mu = 2$ - opposite-charge satisfying the analysis muon selection;
- $p_T(\mu_2) \leq 3.5 \text{ GeV}$ or $\Delta R(\mu_1, \mu_2) < 0.3$. This requirement ensures the analysis is orthogonal, that is, non-overlapping in terms of event content, with the previously published soft lepton analysis [27];
- Event BDT score $\text{BDT} > 0$. This is the main method of selecting signal events while rejecting the SM background. Details are given in Section 4.8.3;
- $\Delta R(\mu_{1,2}, j_1) > 0.4$, where j_1 is the leading jet. The leptons should not be inside the ISR jet;
- ω , ρ^0 and J/ψ invariant mass vetoes. $m_{\ell\ell} \notin [0.75, 0.81] \text{ GeV}$, $m_{\ell\ell} \notin [3, 3.2] \text{ GeV}$.

4.8.2.2 Exclusive track selection

The exclusive track category requires one reconstructed and identified lepton, which can be either an electron or a muon, and an exclusive track, meaning a track that is not identified as a lepton. The track with the highest track BDT score, as described in Section 4.6.6, is picked as the signal lepton candidate. Events in this category must satisfy the preselection and the baseline selection, as well as the following criteria:

- $N_\ell = 1$ lepton passing the analysis muon or electron selection;
- maximum track picking BDT score > 0 , as discussed in 4.6.6;
- event level BDT score > 0 . This is the main method of selecting signal events while rejecting the SM background. , as discussed in Section 4.8.3;
- $\Delta R(\ell, j_1) > 0.4$, where j_1 is the highest-pT jet. The lepton should not be inside the ISR jet.

4.8.3 Binary event classifier

This analysis employs a multivariate classifier to select signal events while optimally rejecting SM background events. The classifier algorithm is a BDT, and its output score is used to define Signal Regions (SRs) as well as Control Regions (CRs). For the dimuon category, one BDT is trained, while for the exclusive track category, a BDT is trained for each lepton flavor and for the two phases of the tracker detector (Phase 0 and Phase 1), resulting in a total of five BDTs.

All BDTs are based on the same architecture, making use of 120 trees with a maximum depth of 3. The BDT training is performed with AdaBoost and GiniIndex separation. The BDTs are trained and evaluated using the TMVA package [35].

For training, signal events are taken from the dedicated samples used to train the track-picking BDT for the exclusive track category, listed in Section 4.5.2 for the signal, and SM samples listed in Section 4.5.1 for the background. For the exclusive track category, MC from 2016 and 2017 are used to represent Phase 0 and Phase 1 of the tracker, respectively. For the dimuon category, only 2017 MC is used to represent both phases, with an added systematic uncertainty resulting from this choice.

For the signal the same broad range of higgsino parameter μ ($\tilde{\chi}_1^\pm$) is used as was considered for the track-picking BDT training sample, but only the range of Δm targeted by the analysis. For Phase 0, Δm^0 is selected in the range of [0.3, 4.3] GeV and μ is selected in the range [100,130] GeV. For Phase 1, Δm^\pm is selected in the range of [0.3-4.6] GeV and μ is selected in the range of [100-500] GeV. The preselection and baseline selection is applied to the events included in the training, as well as a subset of the selection criteria listed in Section 4.8.2.1 and Section 4.8.2.2 as follows:

- $N_\mu = 2(1)$ 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 score > 0 for the exclusive track category.

The training was conducted without using MC weights to avoid possible overtraining issues. This choice does not compromise the performance of the BDT because the kinematics of low- p_T leptons are similar across most SM background production processes. When examining the distributions of input variables in the following sections, this fact must be taken into account. The distributions are plotted 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. Each BDT output working point results in a different signal efficiency depending on the signal parameter values. As will be seen later, one does not use a single value of BDT with a simple cut and count. Instead, the Signal Regions (SRs) are binned according to BDT output values. Therefore, the ROC curve is plotted with a default cut of 0.0 for the sake of completeness. To fully estimate the power of the training, one needs to consider the significance when each signal point has been properly weighted together with the background processes from the SM.

4.8.3.1 Dimuon category

The training samples for the dimuon category contain 4350 signal events and 21842 background events. The BDT evaluated in statistically independent samples of the same size in order to identify any overtraining. The distributions of the testing samples superimposed on the training samples, as well as the ROC curve, are shown in Figure 4.35. No significant overtraining is

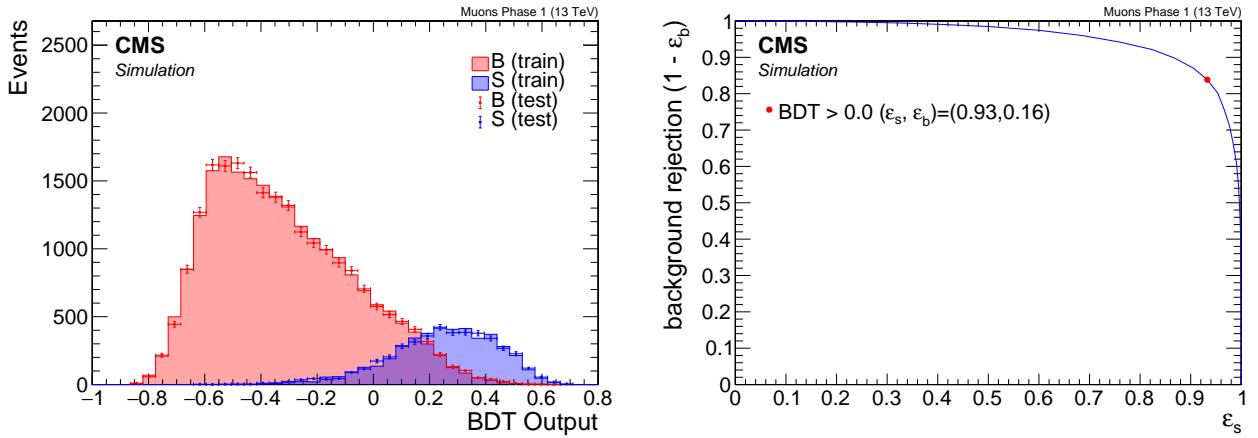


Figure 4.35: Dimuon BDT output (left) and ROC curve (right).

Table 4.9: Dimuon BDT input variables ranked in order of importance, as reported in the TMVA performance summary table.

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	$\min \Delta\phi \left(\vec{H}_T^{\text{miss}}, \vec{j} \right)$	
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	$\left \Delta\phi \left(\ell_2, \vec{H}_T^{\text{miss}} \right) \right $	
13	$\left \Delta\phi \left(\ell_1, \vec{H}_T^{\text{miss}} \right) \right $	
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 $m_{\tau\tau}$

observed. The BDT takes 18 variables as input, listed in Table 4.8.3.1 in decreasing order of importance ranking.

Distributions of the input variables to the BDT are shown in Figure 4.36.

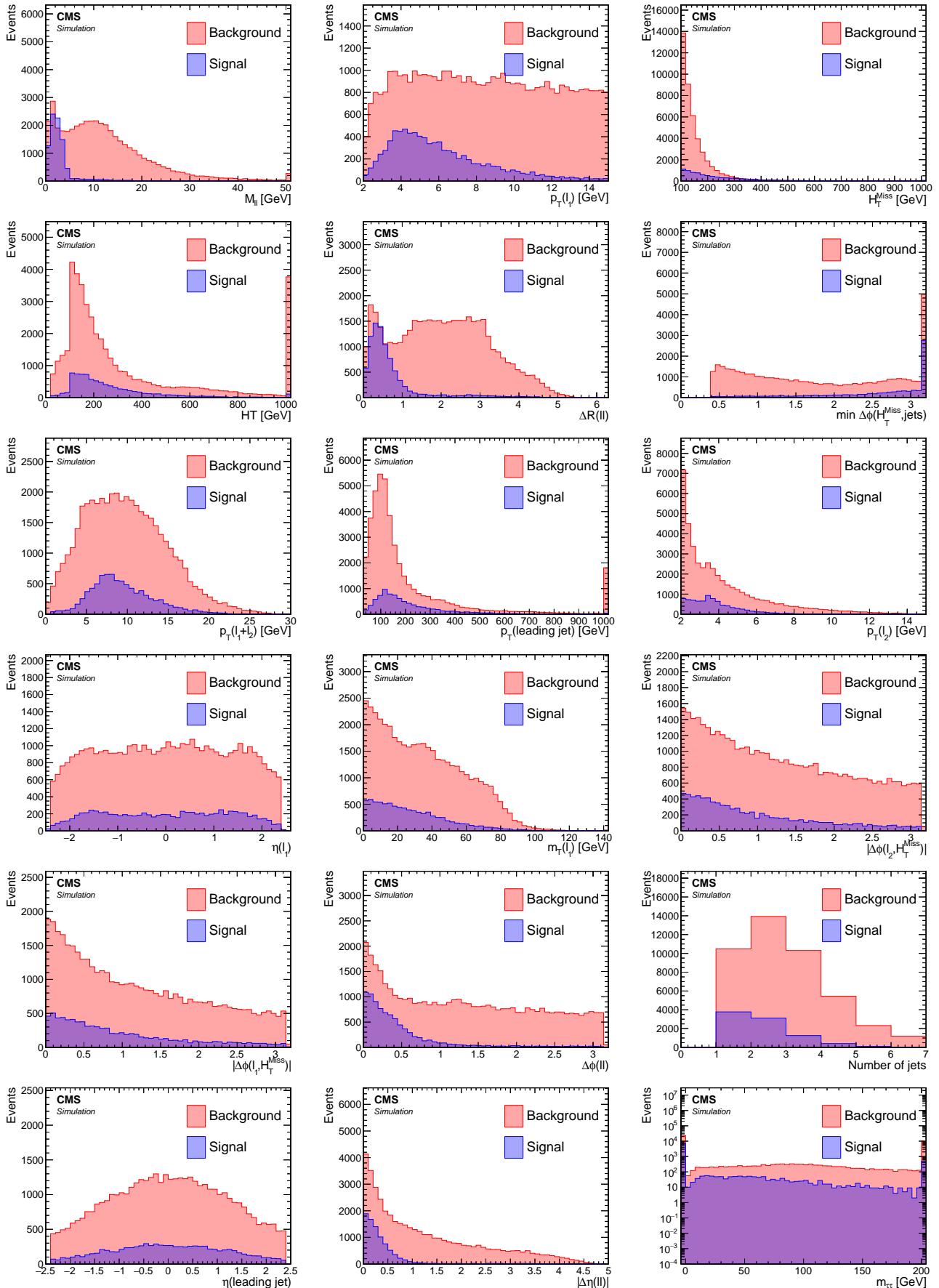


Figure 4.36: Dimuon BDT training input variables. The plots are ordered by importance ranking.

4.8.3.2 Exclusive track category

The training samples for Phase 0 for the exclusive category contain 7863 (1750) signal events and 55765 (29135) background events for muons (electrons). For Phase 1, the exclusive category contain 5266 (1332) signal events and 51308 (31149) background events for muons (electrons). The distributions of the testing samples superimposed on the training samples are shown in Figure 4.37. The ROC curves are seen in Figure 4.38. No over training is observed.

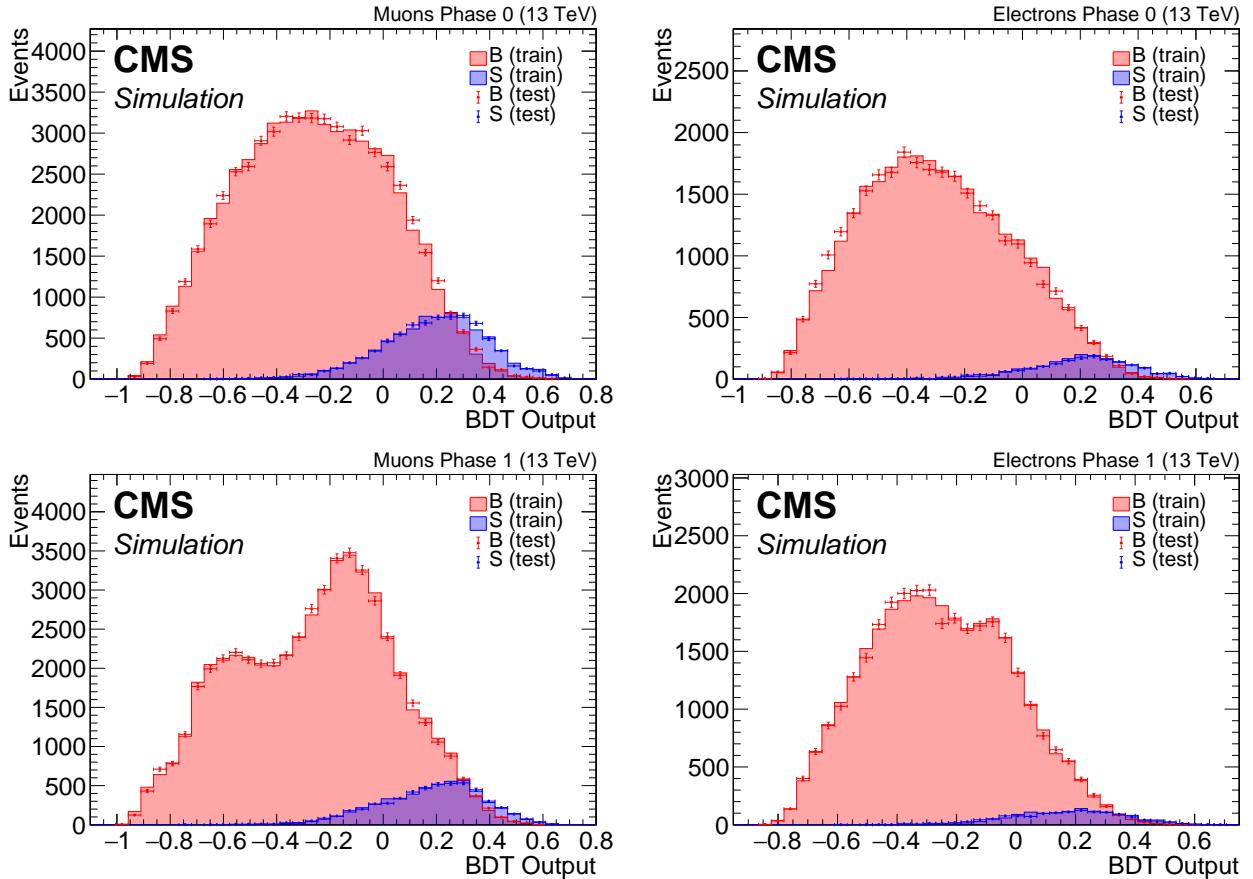


Figure 4.37: Exclusive track category BDT output in Phase 0 (top) and Phase 1 (bottom) for muons (left) and electrons (right).

The training uses 18 different variables listed in Table 4.8.3.2 in decreasing order of importance ranking. Since the ranking is slightly different in the four trainings, the order in the case of the muons of phase 1 is chosen to be listed here. The fully identified lepton is denoted as ℓ and the non-identified lepton track as t .

Distributions of the input variables to the BDT training can be seen in Figure 4.39. 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 fully weighted distributions will be shown in order to asses 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.

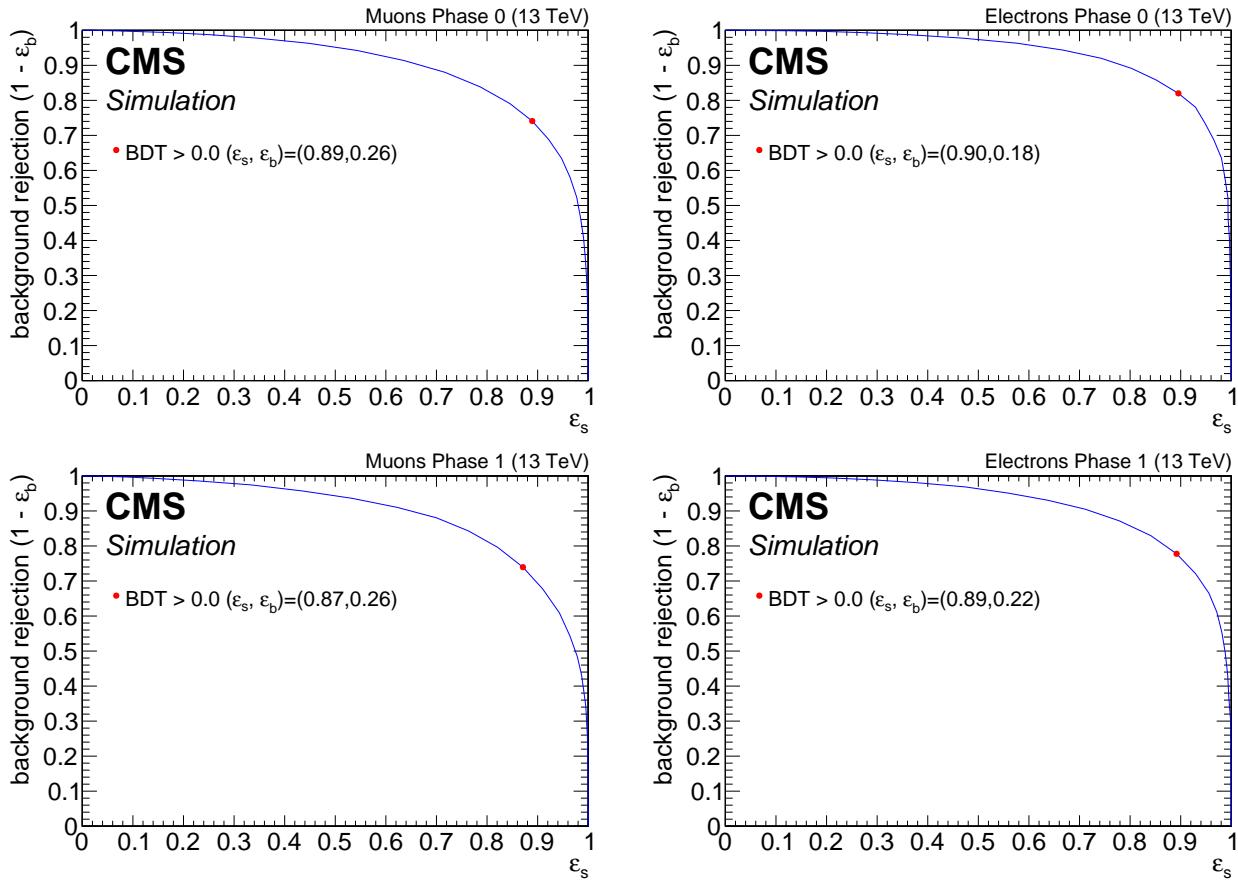


Figure 4.38: Exclusive track category ROC curves in Phase 0 (top) and Phase 1 (bottom) for muons (left) and electrons (right)

Table 4.10: Exclusive track BDT input variables

Rank	Variable	Description
1	$p_T(\ell)$	lepton p_T
2	H_T	
3	H_T^{miss}	
4	$\min \Delta\phi \left(\vec{H}_T^{\text{miss}}, \vec{j} \right)$	
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) $	

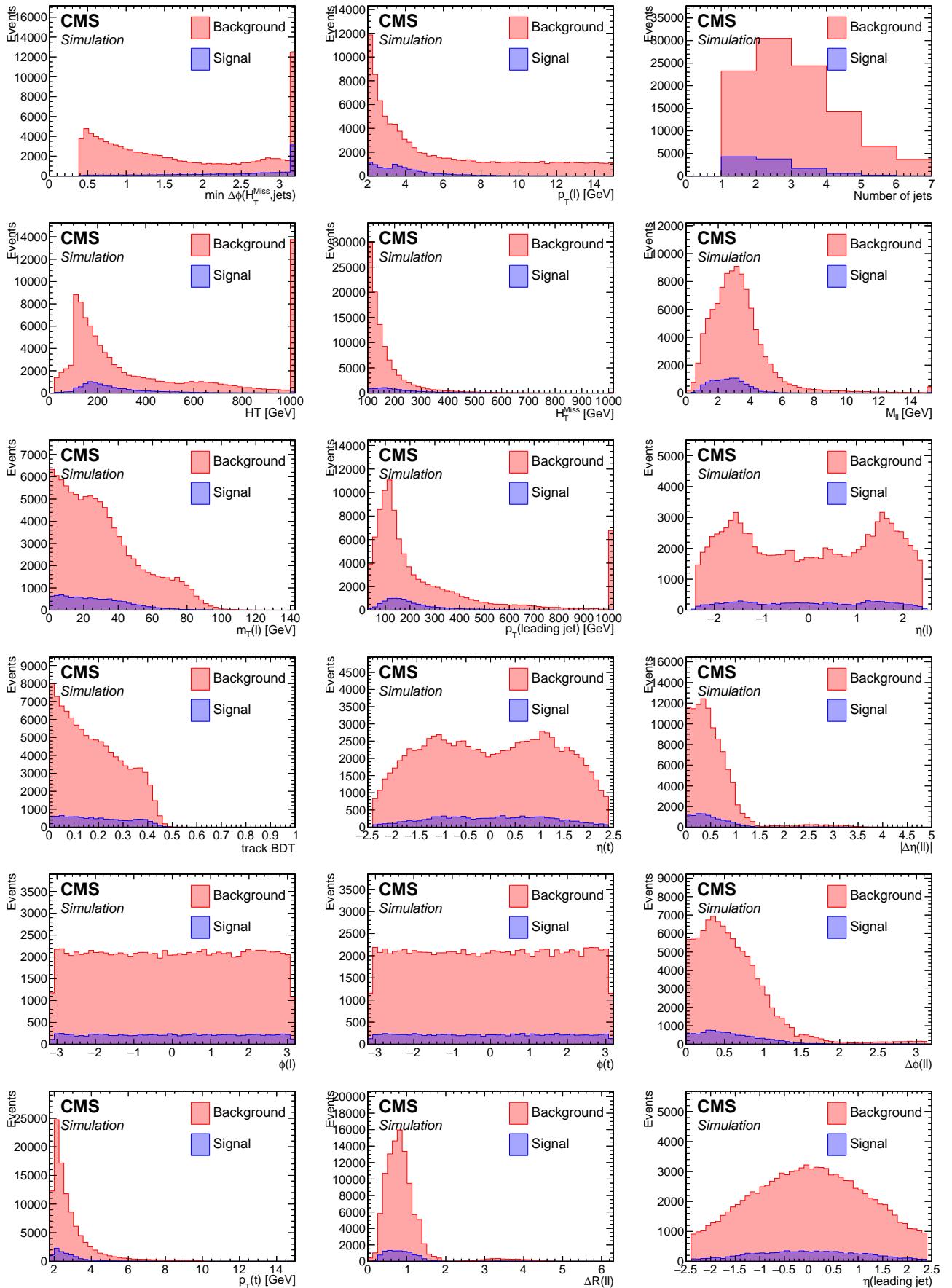


Figure 4.39: Exclusive track BDT training input variables. The plots are ordered by importance ranking.

4.9 Characterization and estimation of the Standard Model backgrounds

Backgrounds arising from SM processes as well as fake or spurious tracks and leptons present a practical challenge for the analysis, given the unique low-momentum phase space of the selected objects. The characterization of the SM backgrounds is examined in Section 4.9.1, while the methods for estimating the background rates in the signal region is described in Section 4.9.2.

4.9.1 Characterization of the Standard Model backgrounds

Processes which contribute to event counts in the signal region, but which are not attributed to the signal process, are referred to as backgrounds. Backgrounds can arise due to SM processes with final states closely resembling the signal, or due to detector effects and mismeasurements. In the current analysis, an example of a background in the dimuon category that arises from truly similar physics is Drell-Yan. In a Drell-Yan process, opposite-charge same-flavor dilepton pairs 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, where one lepton comes from the leptonic decay of the W, and another lepton is due to either mismeasurement, i.e., a fake lepton, or as part of a hadronization process. A comprehensive set of the SM processes has been studied with MC samples, along with descriptions, is given below. The processes are ordered according to their contribution in the SRs of the dimuon category.

- **W in association with jets.** In this SM process, a W boson is produced alongside jets and decays leptonically into a lepton and a neutrino. It can be represented symbolically as $W + \text{jets} \rightarrow \ell\nu$. There are several reasons why this process is a background in this analysis. First, since a neutrino is present in the final state, there can be significant real missing transverse momentum. Second, the very low transverse momentum p_T threshold of the analysis muons allows a considerable rate of either a fake misidentified lepton or a low- p_T lepton originating from a hadronization process to pass the analysis selection.
- **Z in association with jets.** In this SM process, there is a production of a Z boson alongside jets, decaying into two neutrinos. It can be written schematically as $Z + \text{jets} \rightarrow \nu\bar{\nu}$. The two neutrinos in this process contribute to true missing transverse momentum in the event. The lepton and track candidates can either be fake, or come from either a decay of a meson produced in the hadronization process.
- **Drell-Yan process.** DY events occur when a quark from one proton and an antiquark from the other proton annihilate, creating a virtual photon or Z boson that 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, as used 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 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 then become a background to this analysis.
- **Ditop.** When two top quarks are produced, $t\bar{t}$, each top decays to a W boson and a b quark, with a branching fraction close to 100%. The W boson can decay to a charged lepton and a neutrino, contributing to real missing transverse momentum and, given the general abundance of low- p_T tracks and fake leptons, can satisfy the dimuon or track+muon selection. Despite the b-tagged jet veto applied as a component of the baseline selection, a non-negligible rate of $t\bar{t}$ events persists in the signal region.

- **Diboson and rare processes.** In the plots presented in the following section, diboson processes (VV) is distinguished from higher-order productions such as three bosons, which are collectively referred to as *rare*. The ways in which they can be selected in the SRs are similar to the single boson case. However, the higher-order multiplicity events have much lower production cross sections, and are therefore 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} . Most E_T^{miss} present in a QCD event is due to the mismeasurement of jet energy. The relatively high E_T^{miss} cut, in combination with requiring $\min \Delta\phi(\vec{H}_T^{\text{miss}}, \vec{j}) > 0.4$, eliminates almost all QCD background. Accounted for using the jetty-background method in Section 4.9.2.1.
- **Resonances.** Resonances are composite particles, namely mesons or baryons, which can later decay into leptons. The largest contribution in this category comes from the J/ψ , which has a mass of 3.1 GeV and a relatively high cross section. To reduce this background, invariant mass vetoes for the ω , ρ^0 , and J/ψ are applied in the ranges of $m_{\ell\ell} \in [0.75, 0.81]$ GeV and $m_{\ell\ell} \in [3, 3.2]$ GeV.

To gain an understanding of the proportion of each background process, luminosity-weighted MC distributions of key observables are examined in the baseline region, including that of the BDT classifier score as well as of a few important inputs to the BDT. Data taking conditions based on the year 2017 are assumed for this composition study.

4.9.1.1 Dimuon category

The cross-section and luminosity weighted distribution of the BDT output is shown in Figure 4.40 for the dimuon category. Six SR bins are defined in the range of the BDT output score greater than 0, and labeled by numbers ordered in increasing sensitivity. The largest backgrounds in the dimuon channel are $t\bar{t}$, $Z + \text{jets} \rightarrow \nu\nu$, and $W + \text{jets} \rightarrow \ell\nu$, with a small contribution from Drell-Yan processes mainly due to $Z/\gamma^* \rightarrow \tau^-\tau^+$. Figure 4.41 shows the top ten input observables to the BDT ranked by importance for the training.

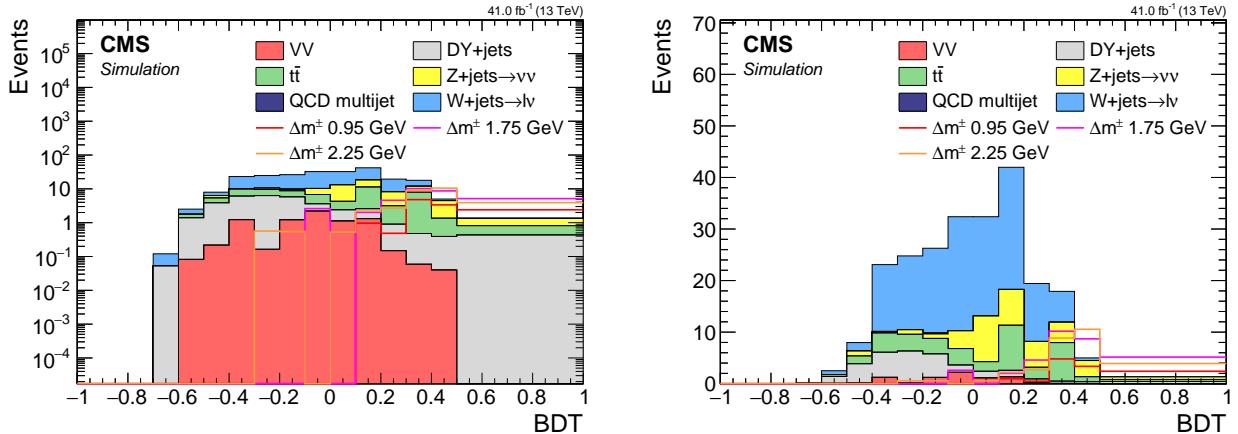


Figure 4.40: Dimuon 2017 simulation BDT score in log scale (left) and linear scale (right).

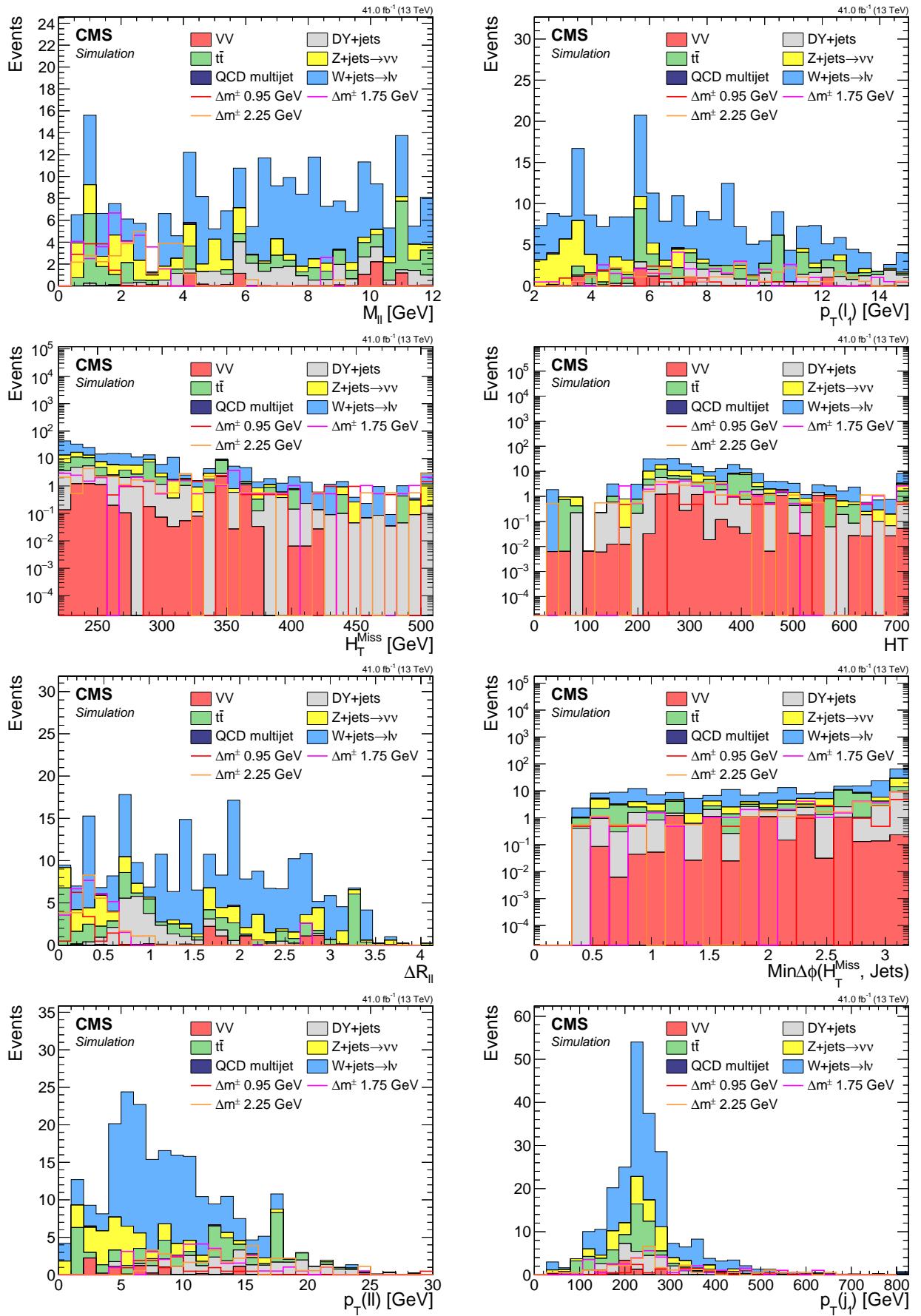


Figure 4.41: Dimuon 2017 simulation BDT inputs for the top 10 ranked observables.

4.9.1.2 Exclusive track category

As described before, there are four BDTs in the exclusive track category, one for each of the two lepton flavors and each of two pixel tracker phases. The distribution of the muon+track category is shown in Figure 4.42. Figure 4.43 shows the top eight input observables to the BDT, ranked by importance for the classifier. It is weighted to 2017 luminosity and uses 2017 simulated data. A few signal points are to indicate the signal-like regions of phase space.

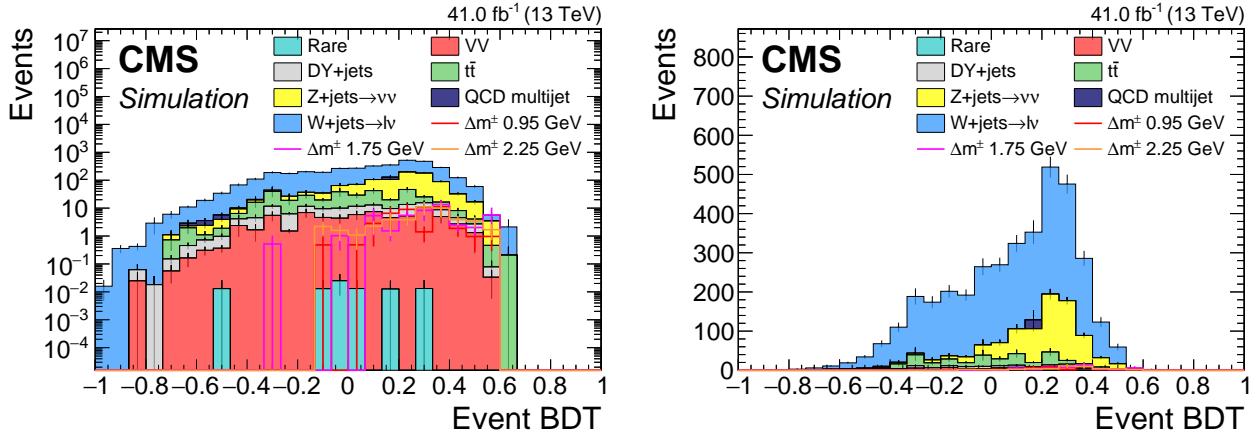


Figure 4.42: track+muon category 2017 simulation BDT output in log scale (left) and linear scale (right).

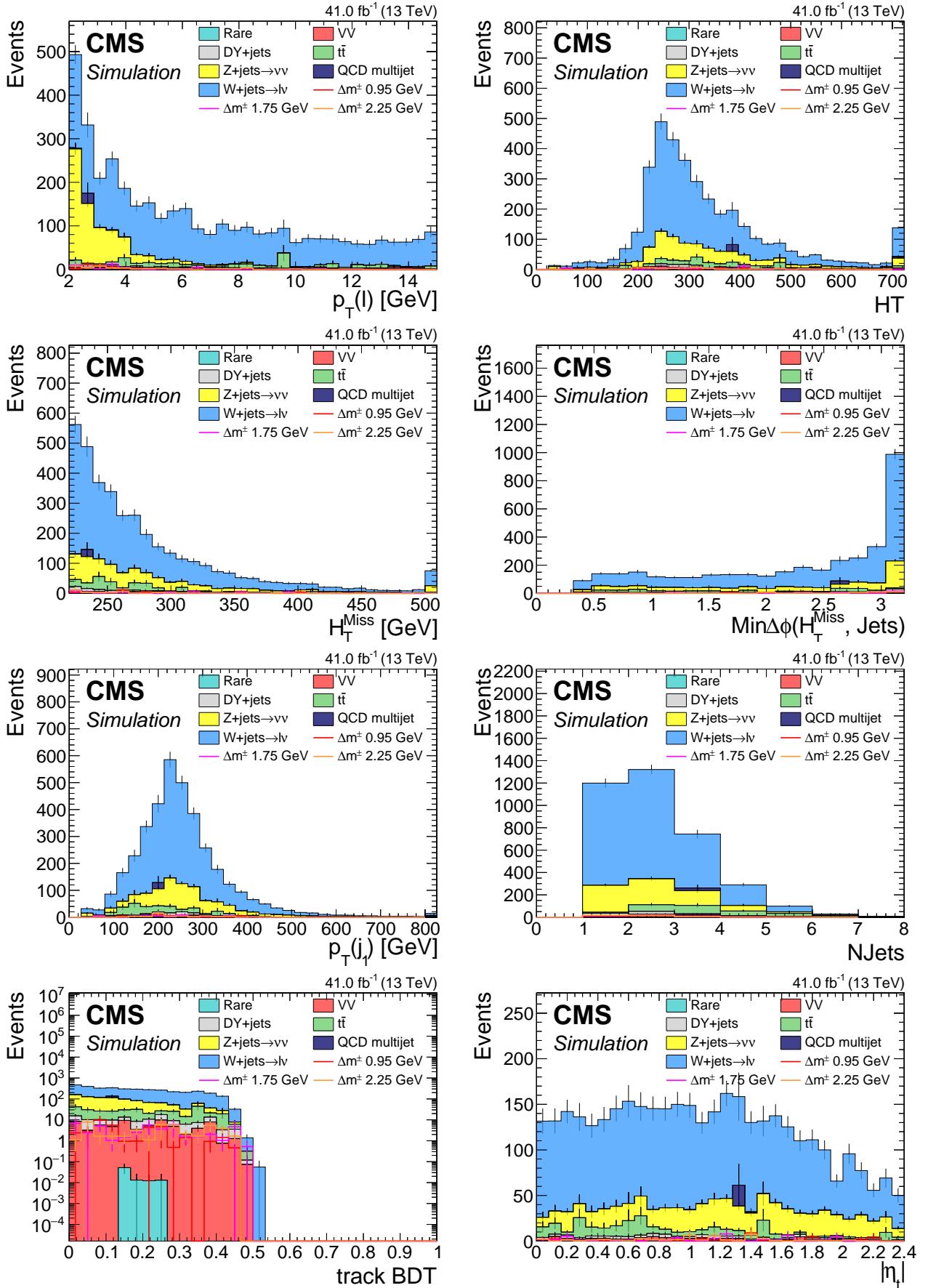


Figure 4.43: Exclusive track plus muon 2017 simulation BDT inputs for the top 8 ranked observables.

4.9.2 Estimation of the Standard Model backgrounds

Accurately predicting the event counts for the Standard Model background is one of the central challenges of the analysis. A widely used method for predicting background counts is MC simulation. MC are weighted to account for production cross-sections and luminosity, and additional correction factors and weights may apply to account for measurement errors, discrepancies between data, and other factors.

Using simulation to estimate the Standard Model background has limitations and disadvantages that can be specific to a given analysis, and depend on the background process under consideration as well as on the observables used in the analysis. The main limitation of simulation is its imperfection. Simulation can never precisely simulate real data due to several factors. Theoretical uncertainties, such as uncertainties on cross sections or branching fractions, can lead to incorrect production rates or normalization. To remedy such effects, simulation is often reweighted using one or more weights derived from a dedicated CR. Another challenging limitation of simulation is its likely misrepresentation of the delicate details of a detector's geometry and response, as well as real-time data-taking conditions which may have varied dynamically throughout a given Run. Some objects and regions of phase space are more prone to discrepancies than others. Using simulation is a reliable method for predicting backgrounds, in which the physics involved has been shown to replicate real data after applying correction factors. In this analysis, the isolated background resulting from the $Z/\gamma^* \rightarrow \tau^-\tau^+$ process is estimated using simulation. However, due to the imperfect modeling of jets in MC, the non-isolated background is modeled using a data-driven method.

A significant challenge arises from the soft nature of the leptons, with low transverse momentum (p_T) and low invariant mass of the order of a few GeV. The sources of background for such events in the standard model include low- p_T resonances produced in hadronization processes, and events where one of the leptons or exclusive tracks is misidentified as one of the signal leptons. These leptons or tracks are often in close proximity to jets in the event. The analysis uses two strategies to estimate this type of background, depending on whether two identified leptons are present, as in the dimuon category, or only one, as in the exclusive track category. The jetty background estimation for the dimuon category is described in Section 4.9.2.1, while the exclusive track background estimation is described in Section 4.9.2.3. As described earlier, a small portion of the background, namely $Z/\gamma^* \rightarrow \tau^-\tau^+$, corresponds to isolated leptons which more closely resemble signal, and the method for estimating this background is described in Section 4.9.2.2.

4.9.2.1 Jetty background estimation

As discussed in Section 4.6, the leptons in the signal are well isolated. The isolation criterion developed for this analysis is the jet-based isolation described in Section 4.6.7. This customized isolation is also a key part of the background estimation, which is described in this section. This background estimation method applies only to the dimuon category, and its estimated contribution is the largest among the two background processes. It is a *data-driven* background estimation method, meaning that the real data, rather than simulation, are used to estimate this background. The name *non-isolated jetty background* refers to the background in which one or both of the leptons are produced in association with jets and are typically in the angular vicinity of a jet. Most of these leptons are rejected by the jet-isolation criteria, but some do manage to pass the isolation if produced far enough from a jet.

This method uses a sideband CR defined by inverting the isolation criteria required for the SR to extract a template that is consistent with the shape of the classifier distribution for the jetty background in the SR. Separate normalization CR, defined in the negative BDT score region, is used to correct for the different production rates of jetty background in the sideband and main band.

The SR is defined by taking BDT output greater than 0, and therefore, by definition, the region with less than 0 becomes a CR. The template extraction region is referred to as the *isolation sideband*. The region defining the SR with the nominal isolation criteria applied is referred to as the isolation *main band*. The SRs are then bins in the isolation *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 the *BDT sideband* or, more elaborately, the *BDT normalization sideband*. Of course though, a *sideband* is still a type of CR.

Lepton candidates in the isolation sideband are by definition within an angular distance ΔR of 0.6 from a lepton-corrected jet. Any jet causing the lepton to fail the jet-based isolation is required to have an original transverse momentum, i.e., transverse momentum before the lepton momentum subtraction, satisfying $15 < p_T < 30 \text{ GeV}$. The upper bound of 30 GeV is chosen because this is the lower bound on the analysis jets, effectively decorrelating the isolation observable from the H_T^{miss} and the number of jets in the event. In the absence of such of an upper bound, a bias in the isolation sideband could, for example, be introduced because requiring a lepton to fail jet-based isolation would require the presence of an additional analysis jet, which is not the case in the main band. The BDT is also not sensitive to these softer jets, and so the shape of the classifier score distribution in the sideband should be unaffected by the isolation requirement, resulting in consistent shapes between the main band and the sideband.

The main assumption underpinning the use of the isolation sideband is that, in the jetty background, the leptons are not isolated but are created in association with jets. Most of them are produced inside the jets, with a distribution that falls off as a function of the angular distance ΔR to jet. By selecting leptons inside the cone around the soft jet, events are picked up that have similar behavior to events where the leptons are outside of those cones. The rate of lepton production inside jets differs from those outside jets, but much about the object and event kinematics is well-matched between the sideband and main band, and only a normalisation correction factor must be applied to bring the two shapes into statistical agreement. The normalisation factor is derived by taking the ratio between the event count in the main band and that in the isolation sideband in the normalisation CR, defined in the region with BDT score less than 0. The event counts in the sideband are then scaled by the normalisation factor to make up the prediction. The prediction in the SR then becomes:

$$\hat{N}_{\text{jetty}}^{\text{SR}} = \frac{N_{\text{main band}}^{\text{norm CR}}}{N_{\text{sideband}}^{\text{norm CR}}} \cdot N_{\text{sideband}}^{\text{SR}}, \quad (4.9)$$

where the transfer factor is:

$$\hat{T}F_{\text{jetty}} = \frac{N_{\text{main band}}^{\text{norm CR}}}{N_{\text{sideband}}^{\text{norm CR}}}. \quad (4.10)$$

The transfer factors are listed in Table 4.12.1.

To test 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, a shape comparison is performed in simulation. This shape comparison, also known as a *closure test*, is carried out by evaluating the consistency of the ratio between the predicted and direct MC values with unity. A normalization factor is computed to correctly normalize the isolation sideband. This is ultimately the same procedure carried out on data to derive the data-driven predictions. This section presents the Phase 1 closure test, carried out using 2017 MC. An additional correction has been carried out in the case of Phase 0, which is discussed in Section 4.12.

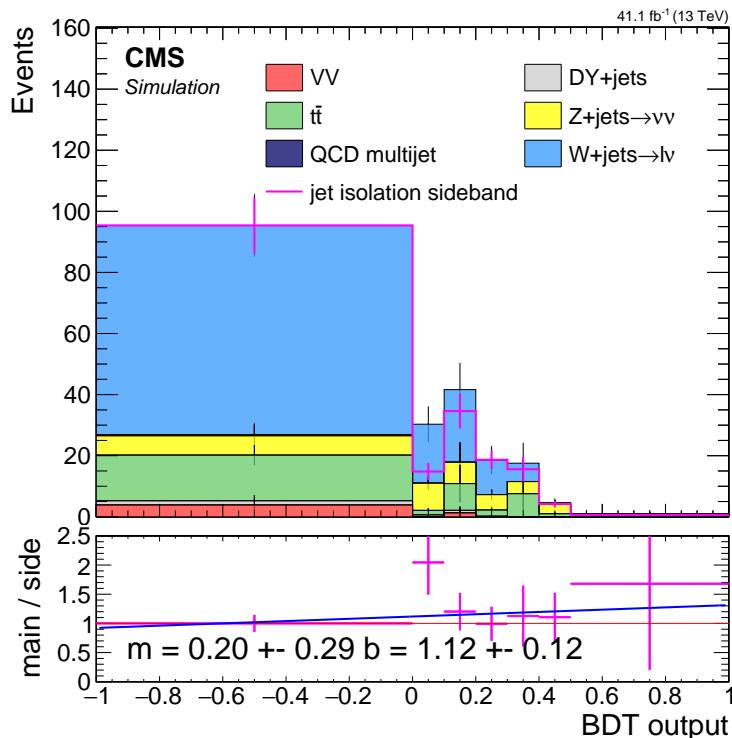


Figure 4.44: Event distributions comprising the Phase 1 jetty background closure test. The stack represents simulation in the isolation main band, $Z/\gamma^* \rightarrow \tau^-\tau^+$ not included, while the pink line represents simulation in the isolation sideband scaled by the normalisation correction factor $\hat{T}F_{\text{jetty}}$. The lower 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 printed.

Figure 4.44 shows the results of the jetty background closure test. The overall shapes are compatible, and the trend line is statistically compatible with a horizontal line at unity, and most bins are statistically consistent with 1. The trend line indicates there is no need for additional correction, but the uncertainty in the trend line constitutes the basis of a systematic uncertainty in the shape of the isolation sideband template. The full list of transfer factors with the associated uncertainties can be found in Section 4.12.1, while the special treatment of the 2016 case is discussed in Section 4.12.2.

4.9.2.2 Ditaū Drell-Yann background estimation

A small amount of background arising from $Z/\gamma^* \rightarrow \tau^-\tau^+$ is also present in the SR, which is the only identified background not accounted for by the jetty method. Since the leptons resulting from 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 is estimated using MC simulation weighted according to a data-to-MC correction factor computed in a dedicated CR that is relatively pure in $Z/\gamma^* \rightarrow \tau^-\tau^+$ background. This control region is constructed by placing requirements on the observable $m_{\tau\tau}$, explained below. If the taus could be fully reconstructed, their system invariant mass $m_{\tau\tau}$ would peak around the Z mass. The Z resonance could then be used as the desired CR rich in ditaū background. However, since leptonic taus are not directly reconstructed, an alternative approach must be formulated.

A widely used method for the reconstruction of the invariant mass $m_{\tau\tau}$ is the *collinear approximation*. First described in [39], it has been used in ATLAS [40] and CMS [41]. In this approximation, it is assumed that each τ produced from Z/γ^* is highly energetic, such that its decay products are collinear, and that the source of missing transverse momentum is the neutrinos. If both τ -leptons are sufficiently boosted, the neutrinos from each τ decay are collinear with the visible lepton momentum. The visible daughter-lepton momentum is used together with \vec{E}_T^{miss} to reconstruct the τ -lepton pair and calculate the invariant mass. Depending on the details of the approximation, one can arrive at a strictly positive distribution for $m_{\tau\tau}$, as in [42], or one that also has negative values as in [43, 44]. The negative values correspond to events where \vec{E}_T^{miss} points more than 90 degrees in ϕ from one of the leptons, which is not consistent with the topology of boosted ditaū events, and thus it is useful to reject negative values in order to purify the CR. The collinear approximation breaks down when the τ s are back-to-back. However, since in the analysis presented in this thesis requires a high- p_T jet and large E_T^{miss} , he considered event topology yields results in sensible values. The signal, as well as other SM processes, are expected to have a smooth and relatively 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 illuminate the logic behind this observable, the following is a derivation of $m_{\tau\tau}$ approximation. The invariant mass is defined as:

$$m_{\tau\tau}^2 = (p_{\tau_1} + p_{\tau_2})^2. \quad (4.11)$$

Assuming that the τ -pair is boosted and the fully leptonic decay products are fully collinear to the τ -leptons, it follows that the transverse momentum of each neutrino pair is proportional to the corresponding τ_i 's transverse momentum by a scale factor ξ_i :

$$\vec{p}_T^{\nu_i} = \xi_i \vec{p}_T^{\tau_i}. \quad (4.12)$$

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

$$\vec{p}_T^{\text{miss}} = \xi_1 \vec{p}_T^{\tau_1} + \xi_2 \vec{p}_T^{\tau_2}. \quad (4.13)$$

Solving the above two equations 4.13 for the two parameters ξ_1 and ξ_2 for each event, the solution becomes:

$$\begin{aligned} \xi_1 &= \frac{\vec{p}_{T_x}^{\text{miss}} \cdot \vec{p}_y^{\ell_2} - \vec{p}_{T_y}^{\text{miss}} \cdot \vec{p}_x^{\ell_2}}{\vec{p}_x^{\ell_1} \cdot \vec{p}_y^{\ell_2} - \vec{p}_x^{\ell_2} \cdot \vec{p}_y^{\ell_1}}, \\ \xi_2 &= \frac{\vec{p}_{T_y}^{\text{miss}} \cdot \vec{p}_x^{\ell_1} - \vec{p}_{T_x}^{\text{miss}} \cdot \vec{p}_y^{\ell_1}}{\vec{p}_x^{\ell_1} \cdot \vec{p}_y^{\ell_2} - \vec{p}_x^{\ell_2} \cdot \vec{p}_y^{\ell_1}}. \end{aligned} \quad (4.14)$$

Equation 4.11 is expanded based on the assumption that the τ 's are boosted and that the four-momenta of the τ s is $p_{\tau_i} = (1 + \xi_i)p_{\ell_i}$:

$$\begin{aligned} m_{\tau\tau}^2 &= (p_{\tau_1} + p_{\tau_2})^2 \\ &= ((1 + \xi_1)p_{\ell_1} + (1 + \xi_2)p_{\ell_2})^2 \\ &= 2m_\tau^2 + 2(1 + \xi_1)(1 + \xi_2)p_{\ell_1} \cdot p_{\ell_2} \\ &\approx 2(1 + \xi_1)(1 + \xi_2)p_{\ell_1} \cdot p_{\ell_2}. \end{aligned} \quad (4.15)$$

This can be negative if one of the ξ_i satisfy $\xi_i < -1$. 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 non-DY processes, such as WW+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

$$m_{\tau\tau} = \text{sign}(m_{\tau\tau}^2) \sqrt{|m_{\tau\tau}^2|}. \quad (4.16)$$

The CR constructed to constrain the $Z/\gamma^* \rightarrow \tau^-\tau^+$ background should have high purity, and thus minimal contamination from SUSY signal and other processes. Figure 4.40 shows that the region of $\text{BDT} < 0$ has negligible signal contamination, and is therefore used as a starting point to build the $\tau\tau$ CR. Figure 4.45 displays the $m_{\tau\tau}$ distributions for the $\tau\tau$ MC in red and the rest of the standard model backgrounds in the stack. The results for the two tracker phases are presented side by side. A clear peak in the $\tau\tau$ background is observed around the mass of the Z boson. A window around the Z boson's mass of [40, 130] GeV is chosen to achieve high purity of about 75% in both phases. Contamination from other backgrounds is removed by first predicting the jetty background count using the data-driven method described in Section 4.9.2.1, and subtracting those counts from the data counts in the $\tau\tau$ dedicated CR. The ratio of data to MC is extracted from this region, with the result 1.2 ± 0.46 (0.29 ± 0.26), which has a relative error of 38% (90%) for Phase 0 (Phase 1).

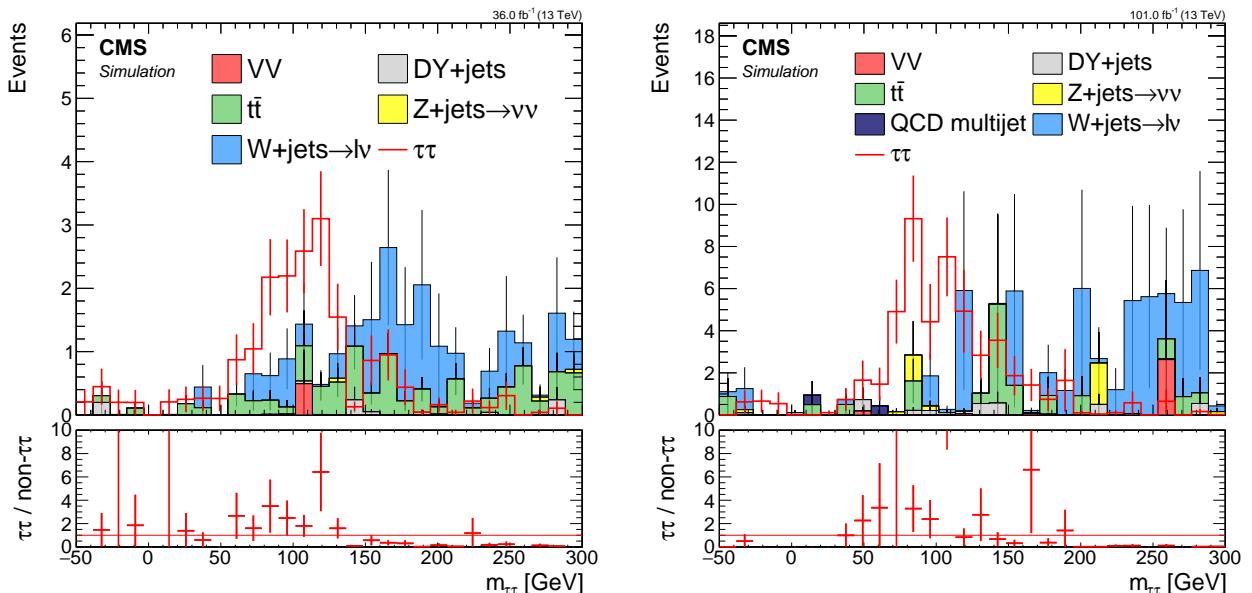


Figure 4.45: Ditàu invariant mass $m_{\tau\tau}$ distributions for phase 0 2016 simulation (left) and phase 1 2017 simulation weighted to luminosity of 2017-2018 data taking period (right). The red line corresponds to $\tau\tau$ simulation, and the stack represents the rest of the standard model background simulation. No overflow bins are plotted in order to clearly show the resonance peak.

4.9.2.3 Exclusive track background estimation

The exclusive track category uses four separate BDTs, one for each lepton flavor, and for each phase. However, the background estimation method is the same for all of them.

The exclusive track category requires one identified lepton according to the selection listed in Sections 4.6.1 and 4.6.2, and one track selected by a procedure described fully in Section 4.6.6. The track is chosen with the highest BDT score among all tracks in each event using the track-picking BDT that was trained to pick up the track that corresponds to the non-identified lepton in the signal event. The chance of selecting a track/lepton pair corresponding to the decay of a single resonant particle is vanishingly small. It is highly likely that the track corresponds to an unrelated charged hadron or is a fake track, meaning a fluke in the tracking pattern recognition procedure.

To devise a reliable background estimation procedure for the exclusive track category, a symmetry is exploited relating to the charge of tracks in the background. The nominal selection requires tracks with opposite charge to the identified lepton, but given that the track is produced independently from the lepton, events with a track of the same charge have otherwise practically indistinguishable characteristics from events with opposite charge pairs. Both the overall rate as well as the shape of the BDT output are generally equivalent, making it an excellent proxy to the true background.

A CR is defined by selecting events with a same-charge lepton-track pair rather than an opposite-charge pair as in the SR. The normalization is fixed by calculating a normalization factor as the ratio between the opposite-charge and 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 using MC data. Figure 4.46 shows the results of the closure tests for muons and electrons for both tracker phases. In each plot, the stack represents SM background for the nominal (opposite-charge) analysis selection lepton-track pair (oc), while the orange line represents the same-charge lepton-track pair (sc). In the ratio panel, which shows the ratio between the opposite-charge to same-charge backgrounds for each bin, the shapes of the nominal and sc backgrounds are seen to be consistent.

After establishing that the method can be used to correctly predict the background, a data-driven normalization factor 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%.

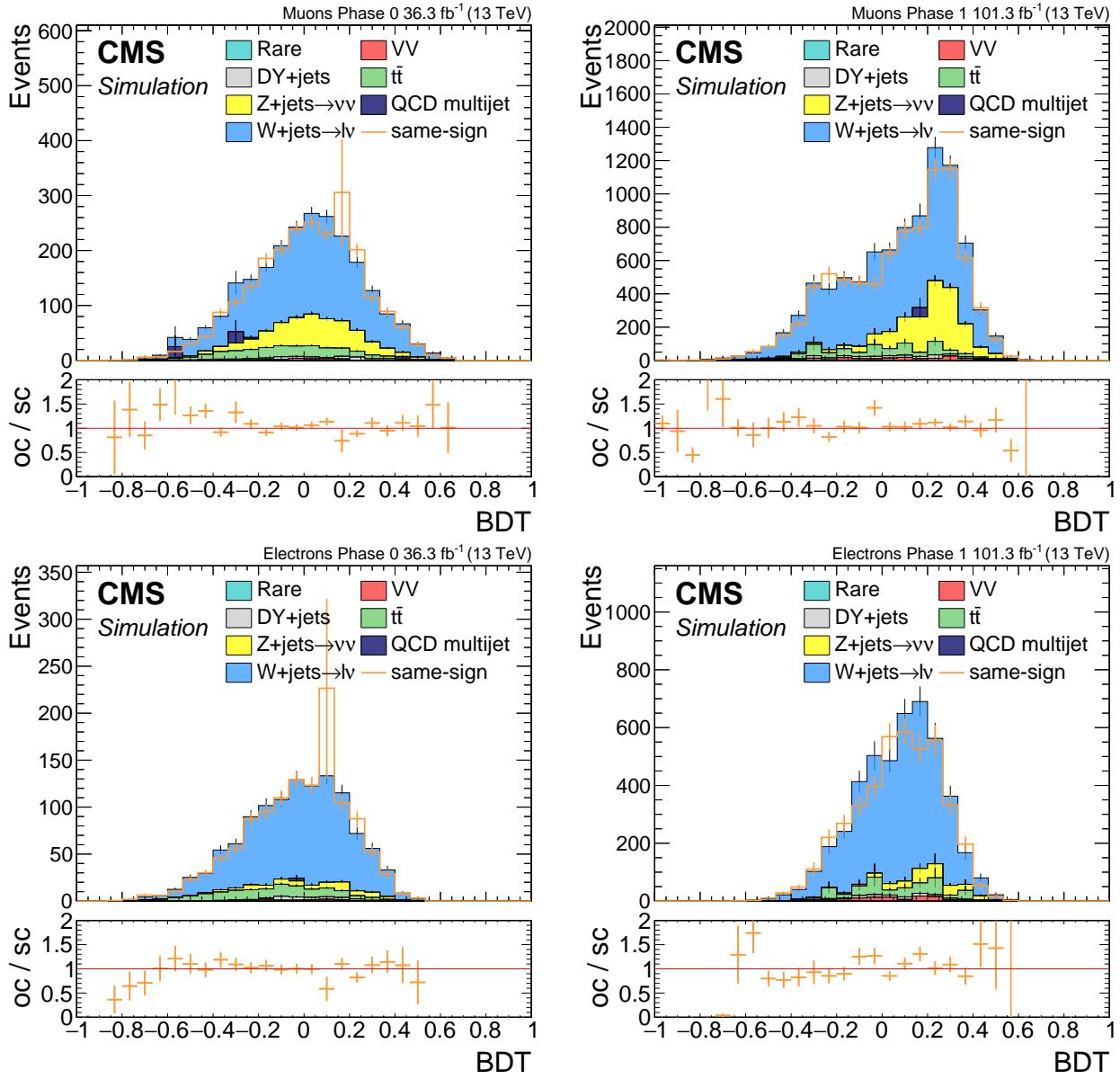


Figure 4.46: Distributions constituting the closure tests for the exclusive track background for the muon+track (top) and electron+track (bottom) for Phase 0 (left) and Phase 1 (right). The stacked histograms represent the SM background for OC pairs, while the orange line is the distribution for SC pairs noramlized according to the method. The lower panel shows the ratio between opposite-charge and same-charge counts for each bin. All uncertainties shown are statistical.

4.10 Data control region plots

MC simulation is used in a number of ways in the analysis, including to train the BDTs and to gain understanding of the composition of the background, and to test the logic of the background methods (closure tests). It is therefore useful to compare distributions of key observables in data and MC to verify that the simulation does not significantly diverge from the data. To avoid unblinding the data in sensitive regions, are made in various CRs known to be devoid of signal.

A useful validation CR is the region obtained by selecting events with an event-based classifier score less than 0. In the following study, this region is examined for the dimuon category. A focus is made on the Phase 1 data set because it is host to the various data quality issues that are further addressed in Section 4.13. The comparison is shown in Figure 4.47. Generally good agreement between data and simulation can be observed in the ratio panel.

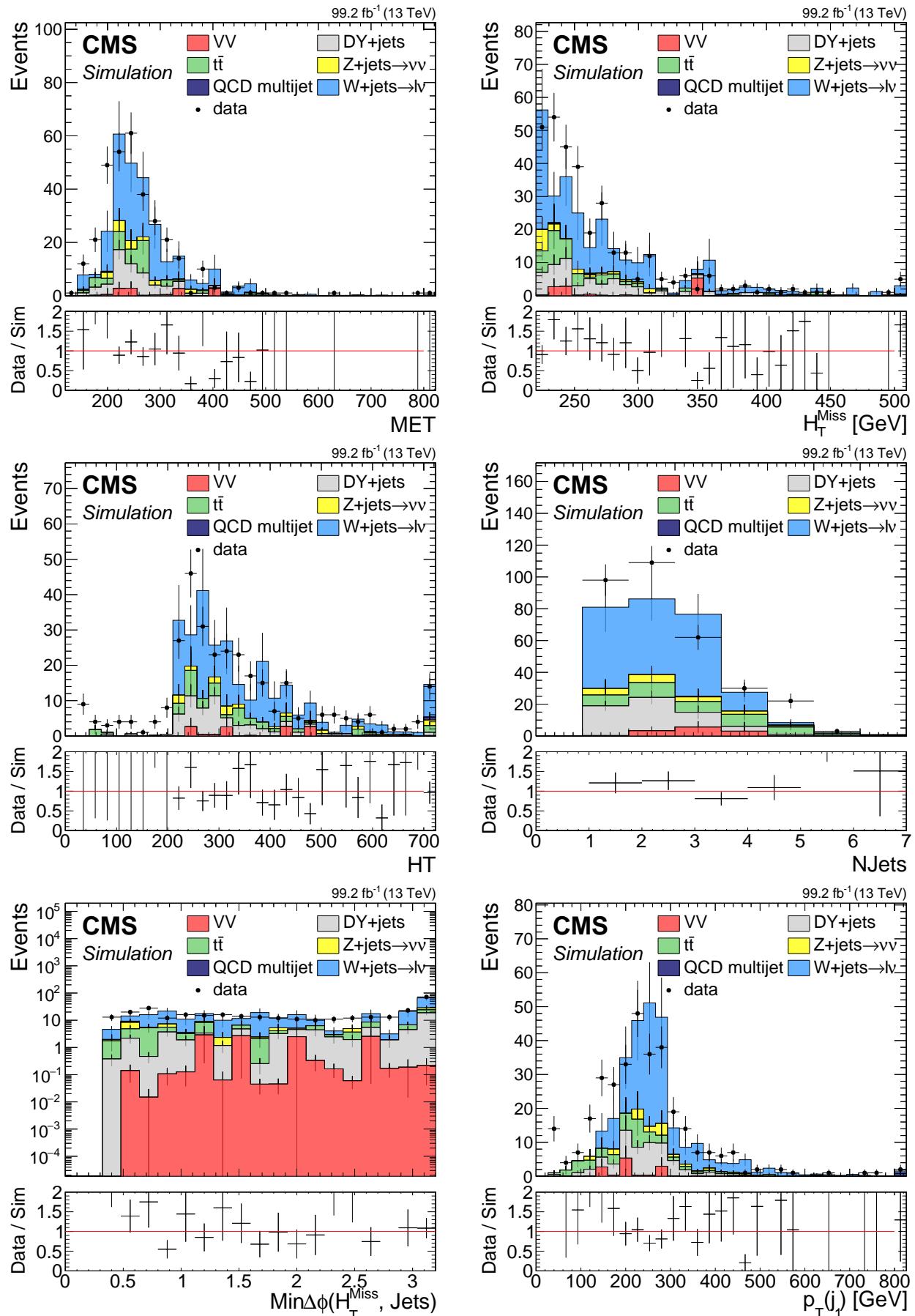
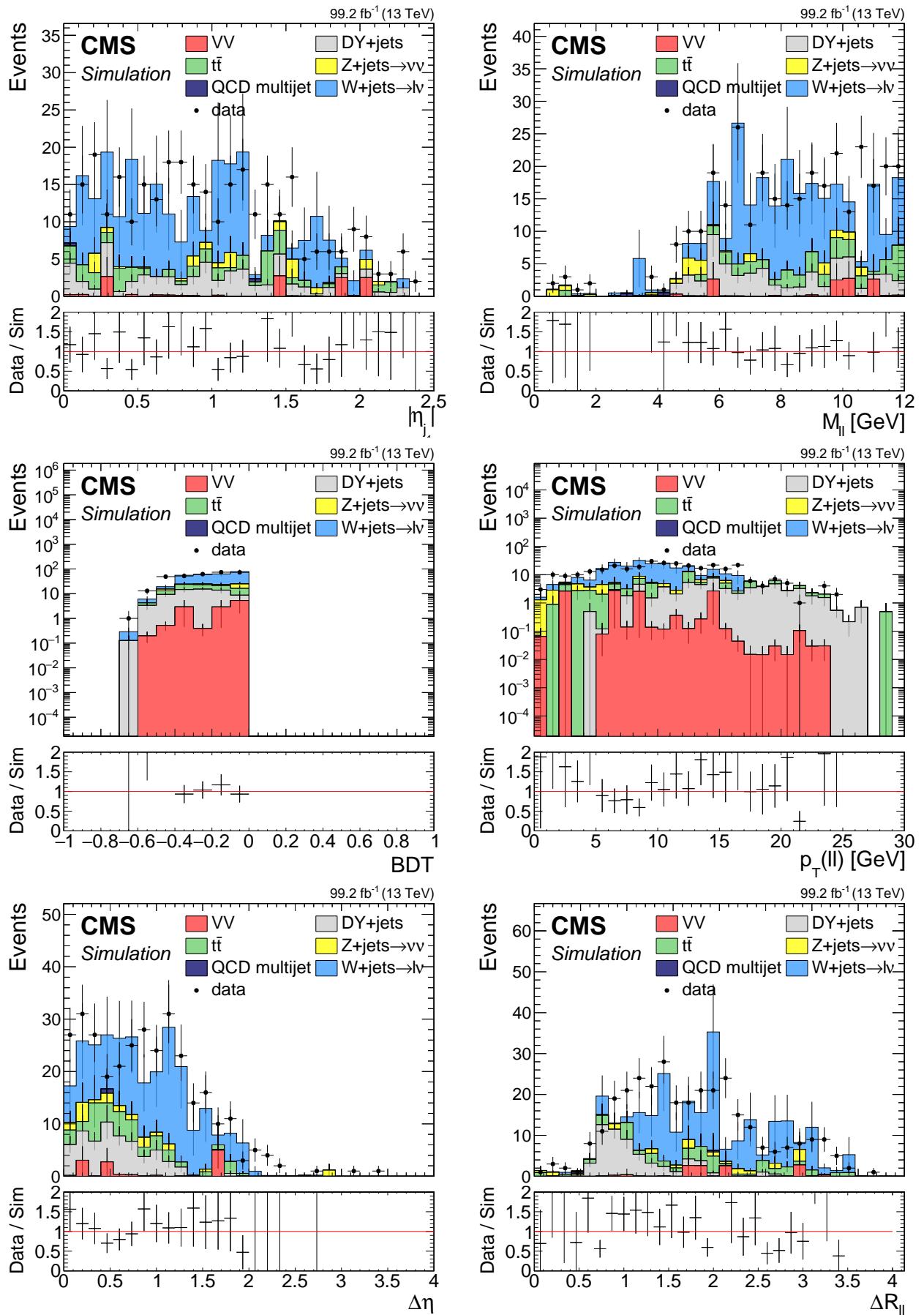


Figure 4.47: Data control region plots for dimuon category in phase 1.



4.11 Optimization of search bins

The signal region is split into various search bins in the range of the event BDT classifier output greater than zero. The final likelihood fit is performed using all of the bins simultaneously, and this approach is a type of *shape analysis*. The significance is computed in each bin, and the individual significance values are then combined to yield single significance value for a given signal hypothesis.

As a general rule of thumb, the signal purity increases as a function of the BDT output score. This means that the most significant bin is likely to be to the right end of the distribution. Finding an ideal choice of bin boundaries can be challenging because the distributions are not smooth, but are made up of event counts with potential statistical fluctuations. The first step in defining the SRs is defining the rightmost division that becomes the left edge of the most sensitive bin, stretching all the way up to the maximum BDT output value of 1. To choose this bin boundary, a scan is performed over all possible lower thresholds on the BDT score in the considered range, employing a step size ε . 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.

Three open points regarding the binning optimisation warrant further elaboration. The first is the choice of measure for estimating the significance. Since the final significance, combination, and exclusion limit are calculated using the CLs method with asymptotic limits, and is somewhat intractable for a study like this [45], a simple estimate is employed at this stage, which is reviewed in [46, 47] and referred to as the Z-value. The Z-value is related to the p -value by specifying the corresponding number of standard deviations in a one-tailed test of a Gaussian (normal) variate:

$$Z = \Phi^{-1}(1 - p) = -\Phi^{-1}(p). \quad (4.17)$$

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

$$Z = \frac{\hat{s}}{\sqrt{\hat{b} + \delta\hat{b}^2}}. \quad (4.18)$$

The background event count is estimated using the data-driven methods described in 4.9.2. They all involve counting events in a sideband and multiplying them by a transfer factor computed in a control region:

$$\hat{b} = N_{\text{sideband}}^{\text{SR}} \cdot \text{TF}, \quad (4.19)$$

where the transfer factor TF is given by

$$\text{TF} = \frac{N_{\text{main band}}^{\text{CR}}}{N_{\text{sideband}}^{\text{CR}}}. \quad (4.20)$$

The error propagation formula yields

$$\left(\frac{\delta\hat{b}}{\hat{b}} \right)^2 = \left(\frac{\delta N_{\text{sideband}}^{\text{SR}}}{N_{\text{sideband}}^{\text{SR}}} \right)^2 + \left(\frac{\delta \text{TF}}{\text{TF}} \right)^2, \quad (4.21)$$

which results in

$$\delta\hat{b}^2 = \hat{b}^2 \left[\left(\frac{\delta N_{\text{sideband}}^{\text{SR}}}{N_{\text{sideband}}^{\text{SR}}} \right)^2 + \left(\frac{\delta \text{TF}}{\text{TF}} \right)^2 \right]. \quad (4.22)$$

The second point that needs to be addressed is the choice of which signal point or points to optimize. Each model point yields a different signal event count \hat{s} and therefore produces

different significance values. To select the optimal bin boundaries, a range of signal model points is considered along the edge of the exclusion limit and thus yield to the strongest limit contours.

The third and final point concerns the choice of step size ε . If ε is too small, there will be steps where no events are encountered in either the signal or the background due to the finite statistics. Therefore, encountered background event causes a discrete jump in the significance, an artificial effect that can lead to overtraining. It will produce meaningfully different results given a statistically independent set of events. To avoid overtraining, a relatively large step size of $\varepsilon = 0.05$ was chosen to balance the need for sufficient statistics for all steps in the scan with the benefits of higher granularity.

After the most significant bin has been fixed, the remaining BDT range from 0 to the low edge of the tightest bin is divided equally in order to increase sensitivity, particularly to models with small Δm . 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 listed in Table 4.11.

Table 4.11: Signal Regions

Category	Flavor	Phase	SR	Signal Regions
Dilepton	Muon	all	6	[0, 0.1, 0.2, 0.3, 0.4, 0.5, 1]
Exclusive Track	Muon	0	13	[0, 0.05, 0.1, 0.15, 0.2, ⋯, 0.5, 0.55, 0.6, 1]
Exclusive Track	Muon	1	12	[0, 0.05, 0.1, 0.15, 0.2, ⋯, 0.5, 0.55, 1]
Exclusive Track	Electron	all	11	[0, 0.05, 0.1, 0.15, 0.2, ⋯, 0.5, 1]

4.12 Systematic uncertainties

The measured and predicted observables have uncertainty associated with them, and this must be taken into account in the interpretation of the data. Sources of uncertainty can be experimental in nature, such as uncertainty in the reconstruction efficiency of muons, or theoretical, such as uncertainty in a cross section. Typically, uncertainty that decreases automatically as the number of events increases is statistical, whereas uncertainty that persists after increasing the statistics is referred to as systematic uncertainty. It could well be that a statistical uncertainty in one study becomes a systematic uncertainty in another.

As mentioned, there are uncertainties associated with theoretical calculations and simulation mismodeling (both for FASTSIM and FULLSIM), among other factors. This analysis follows all the recommendations listed by the CMS SUSY Physics Analysis Group (PAG) [48], which includes the study of muon scale factors described in Section 4.6.3. In this section, only the systematic uncertainties that are unique to this analysis are introduced, aside from the muon scale factors. The systematic uncertainties in this analysis are primarily due to the background estimation methods used.

4.12.1 Data driven transfer factors

Data-driven background estimations are used in both the dimuon category, to estimate the jetty non-isolated background, and in the exclusive track background. They involve computing a transfer factor in a dedicated CR of $\text{BDT} < 0$ and applying it in the SRs. The transfer factors are computed as the ratio between the data counts in the main band and the sideband. In the dimuon category, the sideband is the isolation sideband, as described in Section 4.9.2.1, and for the exclusive track category, the sideband is the same-charge sideband, as described in Section 4.9.2.3. These transfer factors have an associated uncertainty due to the statistics in the CR. Table 4.12.1 lists all transfer factors and their associated uncertainties.

Table 4.12: Transfer factors and their associated uncertainties.

Method	Flavor	Phase	Transfer Factor	Uncertainty	Relative uncertainty
Jetty	Muon	0	0.548	0.078	14.2%
Jetty	Muon	1	0.533	0.039	7.3%
$\tau\tau$	Muon	0	0.518	0.411	79%
$\tau\tau$	Muon	1	0.283	0.26	91.8%
Exclusive Track	Muon	0	1.12	0.044	3.9%
Exclusive Track	Muon	1	1.066	0.024	2.2%
Exclusive Track	Electron	0	1.037	0.05	4.8%
Exclusive Track	Electron	1	1.049	0.03	2.8%

4.12.2 Uncertainty in jetty background template

In the section about the background estimation methods, it is explained that the data-driven methods rely on the assumption that the shape of the background in a sideband is the same as in the main band and, therefore, require only a normalization factor to correctly predict the background. The exclusive track category closure plots in Figure 4.46 show no trend, and neither does the Phase 1 closure plot of the jetty background in Figure 4.44. This is also supported by the line fit performed in the ratio panel, which is statistically consistent with a flat line intersecting 1.

In the dimuon category, only one BDT is trained using 2017 simulation, but evaluated for phase 0 (2016) as well. This introduces a slight trend when a line is fit in the ratio panel of the

closure plot in Figure 4.48. The line fit is then used to introduce weights that are applied in an event-by-event manner with the value of the line for the specific BDT value of the event. On the right side of Figure 4.48, one can see the closure plot after said weights have been applied, and it is clear that the trend has been successfully eliminated.

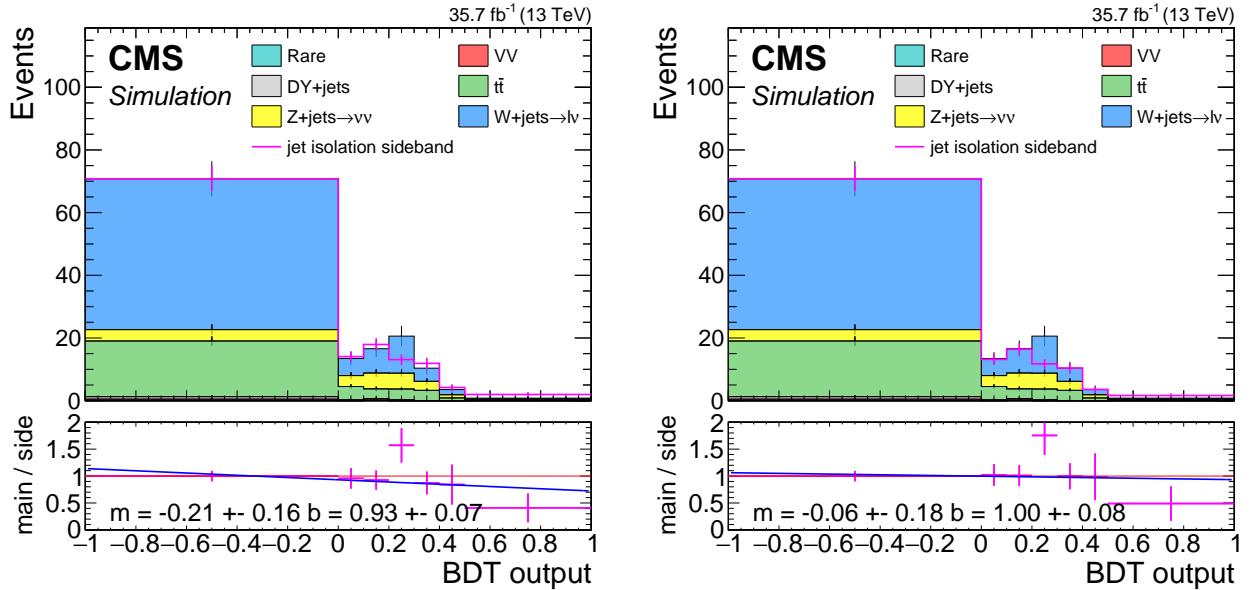


Figure 4.48: Distributions of dimuon 2016 jetty background comprising the closure test with (right) and without (left) fit line weights. The stack represents simulation in the main isolation band excluding $Z/\gamma^* \rightarrow \tau^-\tau^+$, while the pink line represents simulation in the isolation 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 isolatoion 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. In the plot on the right, the line fit weights obtained from the fit on the left plot have been applied.

In addition to the TF uncertainty estimates listed in Table 4.12.1, shape uncertainty based on the line fits are also taken into account. For Phase 1, since the closure plot line fit did not show any trend, the nominal values are taken without applying the line weights. For 2016, the nominal values are taken after the line weights were applied, i.e., from the right plot in Figure 4.48. The alternative prediction, which is fed into the combine tool as the shape systematic uncertainty, is for 2017 the histogram with the line weights applied, and for 2016, since the weights were already applied as the nominal value, the weights of the fit line with the slope varied by 1σ are applied ($m = -0.21 - 0.16 = -0.37$).

4.12.3 Uncertainty in $Z/\gamma^* \rightarrow \tau^-\tau^+$ background

The last background estimation method to consider is the $\tau\tau$, which uses simulation normalized to data in a CR, as explained in Section 4.9.2.2. For background methods that use simulation rather than data, normally a list of uncertainties associated with simulation uncertainties have to be applied. However, as could be seen in Figure 4.49, this background is non-existent in the most sensitive bin, and is very small in the rest of the bins. Therefore, the already very large uncertainties on this background (79%-92%) are dominant enough that all other uncertainties can be safely neglected.

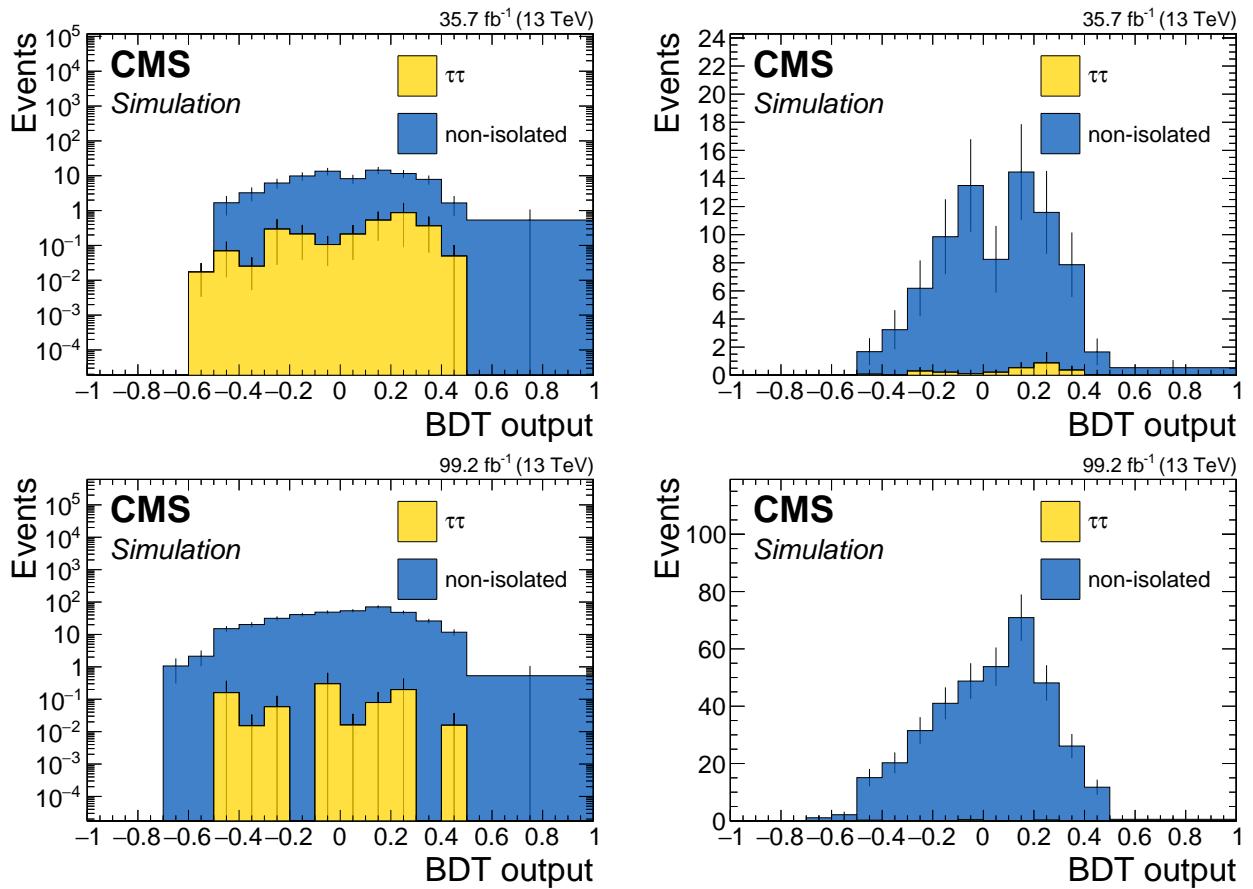


Figure 4.49: Dimuon full background prediction for phase 0 (top) and phase 1 (bottom) both in log scale (left) and linear scale (right). Blue represents the data-driven jetty non-isolated background, while yellow is the $\tau\tau$ background.

4.13 Data quality aspects and background validation

During the Run 2 data taking period, there have been a few detector issues that require some special care. Following the central recommendations, three issues are handled here, namely, L1 prefire rate in 2016 and 2017, ECAL Endcap (EE) noise in 2017, and the HE minus side (HEM) failure in 2018. In the process of dealing with these issues, the jetty background method is also validated in data for selected affected run periods.

4.13.1 L1 prefire issue in 2016 and 2017 data

The L1 prefire issue in 2016 and 2017 occurred due to an ECAL timing error, which was propagated to the L1 trigger primitives. This issue occurred because the trigger system used data from the previous bunch crossing rather than the current one to determine whether an event should be triggered. Events with significant ECAL energy in the region $2.5 < |\eta| < 3$ are affected in 2016 and 2017 data. This can lead to inefficiency and was studied for signal MC samples, as it can potentially lower the signal event count. Prefiring weights were derived and applied to signal and checked against the unweighted events, and no significant effect was observed. Results in the data were also checked with and without the prefiring weights for the most affected period of 2017 by looking at closure plots in a same-charge CR. This serves both to validate that the prefire issue does not affect this analysis and to act as a data validation for the jetty background. Plots can be seen in Section 4.13.4.

4.13.2 EE noise in 2017 data

In 2017 data, an observed excess of fake p_T^{miss} compared to simulation was caused by increased noise in low- p_T jets. Additional noise in the ECAL endcaps in data was identified as the cause of this effect. To deal with this issue, the recommendation is to recalculate p_T^{miss} , excluding jets in the affected phase space. This was done centrally in the process of creating the samples used in this analysis.

4.13.3 HEM failure in 2018 data

Following the power interruptions generated by false fire alarms on Saturday, June 30th, negative endcap Hadron Calorimeter (HCAL) sectors HEM15 and HEM16 could no longer be operated until the end of the 2018 run. The affected $\eta - \phi$ region is $-3.0 < \eta < -1.3$ and $-1.57 < \phi < -0.87$. The first regular physics run affected is 319077. Data and simulation vetoes for objects in the affected region are applied. Same-sign validation plots are made pre-HEM and post-HEM in order to see their effects. The results of these tests are satisfactory and do not imply the need for an additional correction or assessment of additional uncertainty.

4.13.4 Validation in same-sign CR

Figure 4.50 shows the comparison between the predicted and observed background in the same-sign CR. These results serve both as test of the background methods as well as a cross check of the analysis with regards to the data taking issues mentioned above. The results are provided for different data taking periods to check the effects of the data taking issues. The same-sign CR has been selected because it is rich in the primary backgrounds relevant for the search, and because it is devoid of signal events. Overall, good shape agreement is demonstrated between the main band and the isolation sideband.

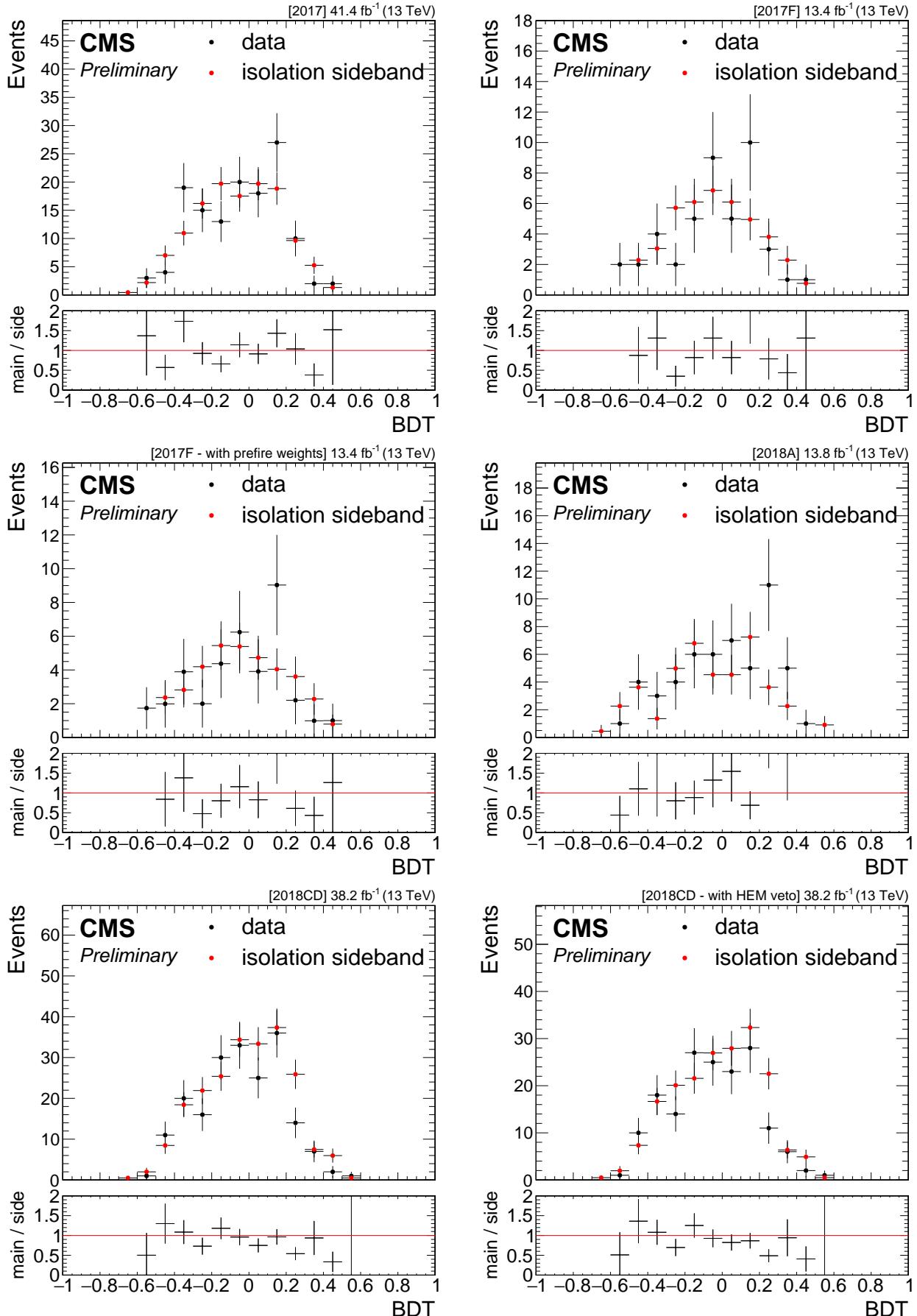


Figure 4.50: Data same sign control validation plots. Black dots show same sign data in the main band, while red dots show same sign data in the isolation side band, normalized in the $\text{BDT} < 0$ region. Ratio panel shows the ratio between them. Going line by line from left to right, the corresponding plots are shown: 2017 data taking period, 2017F data taking period, 2017F data taking period with prefire weights, 2018A data taking period (pre HEM), 2018CD data taking period (post HEM), 2018CD data taking period with HEM veto (post HEM).

4.14 Results and interpretation

Chapter 5

Summary

Chapter 6

Latex stuff

6.1 Some examples

6.1.1 Multiline comment

This is a line in introduction.

6.1.2 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

6.1.3 Tables

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

Table 6.1: 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

6.1.4 Cross References

6.1.4 93 section 6.1.4

6.1.5 Particles

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

6.1.6 Citing

[27] SOS analysis

6.1.7 Glossary

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

6.1.8 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.

6.1.9 fractions

Here's a fraction with a slash:

$$\text{Track Isolation}_\ell = \sum_{\substack{\text{tracks from PV} \\ \text{in } \Delta R < 0.3}} p_T / p_T(\ell) \quad (6.1)$$

6.1.10 Quarantine

From 7.5.1

A Weakly Interacting Massive Particle (WIMP) is a new elementary particle that interacts via gravity and potentially other forces, not part of the SM itself, and is as weak as or weaker than the weak nuclear force, but also non-vanishing in strength. This essentially means that such a candidate is neutral and does not interact via the electromagnetic force.

List of Corrections

Note: fill in signal model stuff	21
Note: make sure we define the different deltaM somewhere	21
Note: This is a fixme note	93
Note: what	93
Note: WHAT THE HELL	93

Glossary

computer is a programmable machine that receives input, stores and manipulates data, and provides output in a useful format. 81

Δm 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., $\Delta m = \Delta m^0 = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$. 24, 28

Δm^0 mass difference between the neutral electroweakinos $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$, i.e., $\Delta m^0 = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$. 24

ΔR separation. 16, 19, 20, 22

η pseudorapidity. 19

E_T^{miss} add description. 11, 12

H_T^{miss} add description. 11, 12, 47

$m_{\ell\ell}$ invariant mass. 16, 18

neutralino add description. 11

ϕ azimuthal angle measured in radians. 19

π ratio of circumference of circle to its diameter. 81

p_T transverse momentum. 11, 15, 20, 24

Acronyms

ATLAS A Toroidal LHC Apparatus. 64

BDT Boosted Decision Tree. 28, 38, 40, 48–50, 52, 56, 57, 59, 66, 68, 73, 74

CERN European Organization for Nuclear Research. 7

CMS Compact Muon Solenoid. 7, 42, 64, 73

CR Control Region. 49, 61, 62, 64–66, 68, 73, 74, 76

CSV Combined Secondary Vertex. 37

DM Dark Matter. 11

ECAL electromagnetic calorimeter. 24, 76

HCAL Hadron Calorimeter. 76

ISR Initial State Radiation. 11, 12, 24, 47, 48

LHC Large Hadron Collider. 7

LSP Lightest Supersymmetric Particle. 11

MC Monte Carlo. 32, 33, 35, 43, 49, 61, 63, 64

PAG Physics Analysis Group. 73

PF Particle Flow. 37, 42

POG Physics Object Group. 32, 35

PU Pile-Up. 37, 42

QCD Quantum Chromodynamics. 11, 12, 47, 56

SF Scale Factor. 32

SM Standard Model. 11, 12, 19, 22, 30, 42, 47–49, 55, 82

SOS Soft-Opposite-Sign. 18–21, 82

SR Signal Region. 49, 55, 56, 62, 66, 71, 73

SUSY Supersymmetry. 11, 73

WIMP Weakly Interacting Massive Particle. 82

WP working point. 24–31, 37, 91

Bibliography

- [1] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory an introduction to quantum field theory*. CRC Press, London, England, Sept., 2019.
- [2] M. Srednicki, *Quantum Field Theory*. Cambridge University Press, Cambridge, England, Jan., 2007.
- [3] R. L. Liboff, *Introductory Quantum Mechanics*. Pearson Education, Philadelphia, PA, 4 ed., Aug., 2002.
- [4] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics, Volume 1*. Quantum Mechanics. John Wiley & Sons, Nashville, TN, June, 1977.
- [5] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics, Volume 2*. Quantum Mechanics. John Wiley & Sons, Nashville, TN, June, 1977.
- [6] P. Dirac, *The principles of quantum mechanics*. International Series of Monographs on Physics. Clarendon Press, Oxford, England, 4 ed., Jan., 1981.
- [7] R. P. Feynman, *Character of physical law*. Modern Library. Random House, New York, NY, June, 1997. <https://www.feynmanlectures.caltech.edu/messenger.html>.
- [8] L. M. Lederman and C. T. Hill, *Symmetry and beautiful universe*. Prometheus Books, Amherst, NY, Oct., 2004.
- [9] “LHC Design Report Vol.1: The LHC Main Ring,” .
- [10] “LHC Design Report. 2. The LHC infrastructure and general services,” .
- [11] “LHC Design Report. 3. The LHC injector chain,” .
- [12] **CMS** Collaboration, S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,” *JINST* **3** (2008) S08004.
- [13] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JINST* **15** (2020) P10017, [arXiv:2006.10165 \[hep-ex\]](https://arxiv.org/abs/2006.10165).
- [14] **CMS** Collaboration, V. Khachatryan *et al.*, “The CMS trigger system,” *JINST* **12** (2017) P01020, [arXiv:1609.02366 \[physics.ins-det\]](https://arxiv.org/abs/1609.02366).
- [15] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Particle-flow reconstruction and global event description with the CMS detector,” *JINST* **12** (2017) P10003, [arXiv:1706.04965 \[physics.ins-det\]](https://arxiv.org/abs/1706.04965).
- [16] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP* **04** (2008) 063, [arXiv:0802.1189 \[hep-ex\]](https://arxiv.org/abs/0802.1189).

- [17] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual,” *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097 \[hep-ph\]](https://arxiv.org/abs/1111.6097).
- [18] CMS Collaboration, CMS Collaboration, “Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid,” CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015. <http://cds.cern.ch/record/2020886>.
- [19] CMS Collaboration, A. M. Sirunyan *et al.*, “Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector,” *JINST* **14** (2019) P07004, [arXiv:1903.06078 \[hep-ex\]](https://arxiv.org/abs/1903.06078).
- [20] CMS Collaboration, S. Chatrchyan *et al.*, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST* **9** (2014) P10009, [arXiv:1405.6569 \[physics.ins-det\]](https://arxiv.org/abs/1405.6569).
- [21] CMS Tracker Group Collaboration, W. Adam *et al.*, “The CMS phase-1 pixel detector upgrade,” *JINST* **16** (2021) P02027, [arXiv:2012.14304 \[physics.ins-det\]](https://arxiv.org/abs/2012.14304).
- [22] CMS Collaboration, CMS Collaboration, “Track impact parameter resolution for the full pseudo rapidity coverage in the 2017 dataset with the CMS phase-1 pixel detector,” CMS Detector Performance Summary CMS-DP-2020-049, 2020. <https://cds.cern.ch/record/2743740>.
- [23] CMS Collaboration, A. M. Sirunyan *et al.*, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JINST* **13** (2018) P06015, [arXiv:1804.04528 \[physics.ins-det\]](https://arxiv.org/abs/1804.04528).
- [24] CMS Collaboration, A. M. Sirunyan *et al.*, “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS,” *Eur. Phys. J. C* **81** (2021) 800, [arXiv:2104.01927 \[hep-ex\]](https://arxiv.org/abs/2104.01927).
- [25] CMS Collaboration, CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018. <https://cds.cern.ch/record/2621960/>.
- [26] CMS Collaboration, CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019. <https://cds.cern.ch/record/2676164/>.
- [27] 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 \[hep-ex\]](https://arxiv.org/abs/2111.06296).
- [28] https://gitlab.cern.ch/cms-muonPOG/MuonReferenceEfficiencies/-/tree/master/EfficienciesStudies/2016_legacy_rereco/Jpsi.
- [29] <https://indico.cern.ch/event/791428/contributions/3287172/attachments/1781977/2899551/2016LegacyReReco2.pdf>.
- [30] BRIL Group, “Bril work suite.” <http://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html>, 2017.
- [31] https://lguzzi.web.cern.ch/lguzzi/Tau3Mu/muonPOG_mediumID_dR.pdf.

- [32] 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>.
- [33] 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>.
- [34] CMS Collaboration, A. M. Sirunyan *et al.*, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV,” *JINST* **13** (2018) P05011, [arXiv:1712.07158 \[physics.ins-det\]](https://arxiv.org/abs/1712.07158).
- [35] 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>.
- [36] <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>.
- [37] 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>.
- [38] 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>.
- [39] 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>.
- [40] ATLAS Collaboration, G. Aad *et al.*, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics,” [arXiv:0901.0512 \[hep-ex\]](https://arxiv.org/abs/0901.0512).
- [41] CMS Collaboration, G. L. Bayatian *et al.*, “CMS technical design report, volume II: Physics performance,” *J. Phys. G* **34** no. 6, (2007) 995–1579.
- [42] 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%2Fphysrevd.89.075007>.
- [43] 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%2Fphysrevd.90.115007>.
- [44] 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%2Fjhep04%282015%29147>.
- [45] <https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/>.

- [46] 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>.
- [47] 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].
- [48] <https://twiki.cern.ch/twiki/bin/view/CMS/SUSRecommendationsRun2LegacyTable>.

List of Figures

2.1	Electrons to Muons Feynman diagram	12
3.1	LHC overview	18
3.2	A cutaway diagram of the CMS detector	19
4.1	Signal Models Feynman Diagrams	21
4.2	Signal E_T^{miss} and H_T^{miss} distributions	24
4.3	Signal <i>number of jets</i> and <i>leading jet p_T</i> distributions	25
4.4	Signal <i>number of b-tagged jets</i> distributions	25
4.5	Signal $\min \Delta\phi(\vec{E}_T^{\text{miss}}, \vec{j})$ and $\min \Delta\phi(\vec{H}_T^{\text{miss}}, \vec{j})$ distributions	26
4.6	Signal p_T distributions	27
4.7	Signal p_T distributions split into barrel and endcaps	28
4.8	Signal $ \eta $ distributions	29
4.9	Signal generator level $m_{\ell\ell}$ distributions	30
4.10	Signal $m_{\mu\mu}$ vs. p_T	31
4.11	Distributions of reconstructed $m_{\mu\mu}$ in signal events	32
4.12	Signal generator level ΔR distributions	32
4.13	Event distributions in the plan of $\Delta R(\mu\mu)$ vs. p_T for signal models	33
4.14	Distributions of the reconstructed $\Delta R(\mu\mu)$ in signal events	33
4.15	Angular seperation between reconstructed electrons and the leading jet $\Delta R(j_1, e)$.	37
4.16	Distribution of reconstructed electron p_T with loose ID	37
4.17	Distributions of $ \eta $ of reconstructed electrons with loose ID	37
4.18	Medium and Tight ID WPs distribution of reconstructed electrons	38
4.19	standard isolation and jet-isolation distribution of reconstructed electrons	39
4.20	$ \eta $ distribution of reconstructed electrons with loose ID passing jet-isolation . .	39
4.21	Angular separation between reconstructed muons and the leading jet $\Delta R(j_1, \mu)$.	40
4.22	Distibution in signal events of the p_T of reconstructed muons	41
4.23	Distibution in signal events of the $ \eta $ of reconstructed muons with loose ID before and after $p_T > 2 \text{ GeV}$ cut	41
4.24	Medium ID WP distribution of reconstructed muons	42
4.25	Tight ID WP distribution of reconstructed muons	42
4.26	Distributions of the jet-based lepton isolation of reconstructed muons	43
4.27	Simluation barrel muons fits	45
4.28	Simluation endcaps muons fits	46
4.29	Data barrel muons fits	46
4.30	Data endcaps muons fits	47
4.31	Efficienciy values and scale factors	48
4.32	Muon track BDT training inputs	51
4.33	Track BDT output plots	52
4.34	Track BDT ROC curves	53

4.35 Dimuon BDT output and ROC curve	62
4.36 dimuon training input distribution	63
4.37 Exclusive track category BDT outputs	64
4.38 Exclusive track category ROC curve	65
4.39 exclusive track training input distribution	66
4.40 Dimuon simulation BDT output	69
4.41 Dimuon simulation BDT inputs	70
4.42 Exclusive track plus muon 2017 simulation BDT output	71
4.43 Exclusive track plus muon simulation BDT inputs	72
4.44 Event distributions comprising the Phase 1 jetty background closure test.	75
4.45 Ditau invariant mass distibutions	77
4.46 Exclusive track category closure tests	79
4.47 Data control region plots for dimuon category in phase 1	81
4.48 Dimuon 2016 jetty background closure plots with and without fit line weights . .	86
4.49 Dimuon full background prediction for both phases	87
4.50 Data same sign control validation plots	89

List of Tables

4.1	The preselection criteria, which are applied to all analysis categories.	25
4.2	Generator level efficiency on muons selections	30
4.3	Selection criteria for tags and probes	45
4.4	Input variables to the in-signal track selecting classifier.	51
4.5	Signal efficiency for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.	55
4.6	Background efficiency for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.	56
4.7	Transfer factor for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.	56
4.8	Significance $s/\sqrt{b + \epsilon_b^2}$ for the jet-based isolation scan for the dimuon channel, based on 2016 MC samples.	57
4.9	Dimuon BDT input variables ranked in order of importance, as reported in the TMVA performance summary table.	62
4.10	Exclusive track BDT input variables	65
4.11	Signal Regions	84
4.12	Transfer factors and their associated uncertainties.	85
6.1	Table captions are above the table whereas figure captions are below.	93