DIFFERENTIAL JET PRODUCTION CROSS SECTION MEASUREMENT IN Z + JET EVENTS FROM PROTON - PROTON COLLISIONS AT $\sqrt{S}$ = 13 TEV USING THE CMS DETECTOR AT LHC

by

Ashley Marie Parker

January 2020

A dissertation submitted to the
faculty of the Graduate School of
the University at Buffalo, The State University of New York
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Physics

Contents

List of Tables

List of Figures

Abstract

The standard model of particle physics, while describing our universe well on many scales, has yet to be precisely measured in all energy regimes. Recent theoretical advances in higher order QCD calculations have provided a way to compare the standard model's predictions to precision measurements of data and monte carlo simulation. Within this dissertation, I present a measurement of the double differential jet production cross section as a function of the jet mass and transverse momentum, in events with a Z + Jet topology, with and without a jet grooming algorithm applied. Studying Z + jet events will yield a light quark enriched jet sample, which has not yet been studied at $\sqrt{s}$ = 13 TeV. Comparing groomed and ungroomed jets will allow us the better understand the jet mass in all energy regimes since the groomed jets will have varying amounts of soft and collinear radiation with respect to the ungroomed counterpart.For ungroomed jets, leading-order and next-to-leading order QCD Monte Carlo programs are found to predict the jet mass spectrum in the data reasonably well, with some disagreement at very low and very high masses. For groomed jets, the agreement between the Monte Carlo programs and the data improves overall, and extends lower in jet mass due to the removal of soft and colinear portions of the jet.First-principles theoretical calculations of the groomed jet mass are also compared to the data, and agree with the data within the range of acceptability

of the calculations. Ultimately these measurements will be used to tune Monte Carlo generators, producing more accurate parton showering simulations, leading to tighter constraint of backgrounds in future searches for new physics.

Chapter 1

Introduction

This thesis presents a measurement of the double differential jet production cross section as a function of the jet mass and transverse momentum, in events with a Z + Jet topology, using 136 fb^{-1} of data acquired from the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) during Run 2 of data taking. This measurement improves upon the previous measurement by CMS experiment at 7 TeV center-of-mass energy [1] and the analogous dijet measurement at 13 TeV [2] . There is a similar measurement by ATLAS experiment in the dijet channel [3] however this is the first measurement in the Z + Jets channel to be presented at 13 TeV.

The standard model of particle physics, while describing our universe well on many scales, has yet to be precisely measured in all energy regimes. Recent theoretical advances in higher order QCD calculations have provided a way to compare the standard model's predictions to precision measurements of data and monte carlo simulation. Studying Z + jet events will yield a light quark enriched jet sample, which has not yet been studied at $\sqrt{s}$ = 13 TeV. Comparing groomed and ungroomed jets will allow us the better understand the jet mass in all energy regimes since the groomed jets will have varying amounts of soft and collinear radiation with respect to the ungroomed counterpart.For ungroomed jets, leading-order and next-to-leading order QCD Monte Carlo programs are found to predict the jet mass spectrum in the data reasonably well, with some disagreement at very low and very high masses. For groomed jets, the agreement between the Monte Carlo programs and the data improves overall, and extends lower in jet mass due to the removal of soft and colinear portions of the jet.First-principles theoretical calculations of the groomed jet mass are also compared to the data, and agree with the data within the range of acceptability of the calculations. Ultimately these measurements will be used to tune Monte Carlo generators, producing more accurate parton showering simulations, leading to tighter constraint of backgrounds in future searches for new physics.

Chapter 2

Theoretical Framework

2.1 Introduction To The Standard Model

”Theorists can be wrong; only nature is always right” - David Gross

The Standard Model (SM) of particle physics is a quantum field theory (QFT) description of the strong, weak and electromagnetic forces of nature. The known particles of the SM are; 1 scalar Higgs boson,4 guage bosons, 6 types of quark, and 6 types of lepton.

The quarks and leptons, particles which constitute matter, are fermions, obeying Fermi-Dirac statistics due to their half-integer spin. In contrast,the bosons have integer spin and obey Bose-Einstein statistics. Gauge bosons mediate the 3 fundamental forces and the Higg’s boson is responsible for the electro-weak symmetry breaking which gives mass to the other particles[4].

The fermions are arranged into 3 generations, arranged in columns from left to right on [**?**]

Quantum Chromodynamics, QCD, is the theory of the strong interaction which governs the interactions of quarks and gluons[4].

The SM constitues humanity's most rigorous theory of our universe, providing predictions of observables which have since been measured, in the case of Quantum Electrodynamics, QED, to the highest precision of any scientific theory. Despite the impressive predictions,the gravitational force and more subtle phenomena, such as flavor oscillation of neutrinos [5], indicate the existence of physics beyond the standard model, BSM.

Various attempts have been made to unify the fundamental forces under one theory, thusfar the electromagnetic and weak interactions have been united by electro-weak theory.

The Standard electroweak model can be described $SU(2)xU(1)$ mathematically.

The $SU(2)xU(1)$ guage group is a convolution ($< -$That is not the right word...) of the special unitary symmetry group $SU(2)$ describing 3 mixed massive vector bosons, W_{-} W_{+} Z_0, carriers of the weak nuclear force and the unitary gauge group $U(1)$, describing the lonely massless chargeless photon, of the electromagnetic interaction.

The standard model of the strong interaction is known as quatum chromodynamics, QCD, a non-Abelian guage theory described by the special unitary group $(SU(3)_f)$, where the flavours of quark are the physical manifestation of the symmetry group. This force is mediated by the 8 massless gluons which carry color charge, making QCD more complicated mathematically than QED.

The SM also contains a Higgs boson, an excitation of a scalar Higg's field, which gives rise to spantaneous symmetry breaking of the electroweak theory, providing the particles with mass, but I won't get into that.

The quarks and leptons are arranged in generations according to their relative masses, as shown in Figure 2.2. The table also shows the spins of the particles, the leptons and quarks have half-integer spin, fermions, that obey the fermi exclusion principle, conversely the bosons have half integer spin and therefore obey bose-einstein statitics. Through the SM we interpret the observed hadronic particles, mesons (baryons) , as 2 quark (3 quark) bound states. The existence of spin $\frac{3}{2}$ baryons, which are symmetric bound states in space, spin and flavour and the need to obey Fermi-Dirac statistics, by maintaining total assymmetry of the wavefunction,implies there is another degree of freedom, called color, so that each quark is either red, green or blue. Granted only color singlet, containing either all 3 or 1 and it's anti color, states exist. Furthermore there exists a property of asymptotic freedom where the QCD coupling between quarks and gluons increases as they asymptotically approach one another. There exists a wealth of experiemental data to support the concept of asymptotic freedom despite the fact that rigorous mathematical proof of the exlusion of free quark and gluon states has yet to be acheived.

Assymptotic freedom is a useful property as it allows for perturbative calculations of QCD observables, this is discussed in section XXX.

Nuclei in ordinary matter are composed solely of 1^{st} generation particles, up and down quarks, bound by gluons. Neutral atoms contain an equal number of protons (composed of 2 up quarks and a down quark) and electrons, 1^{st} generation leptons. The main distinction between leptons and quarks, both fermions (particles of $\frac{1}{2}$ integer spin), being that leptons do not experience the color interaction $(SU(3)_f)$ like their quark friends. In each generation there is a quark with charge $Q = +\frac{2}{3}$ (up, charm, top) and another of charge $Q = -\frac{1}{3}$ (down, strange, bottom).

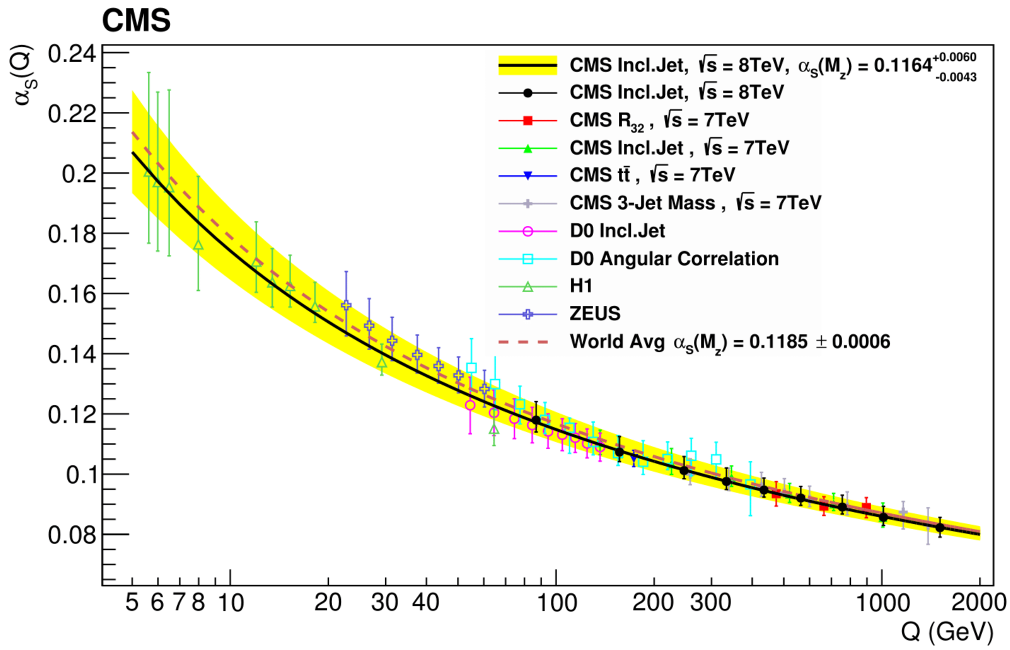

Figure 2.1: The running of the strong coupling constand as compiled by CMS including measurements from CMS and HERA among others [6].

2.1.1 Quantum Electrodynamics

2.1.2 Quantum Chromodynamics

lagrangian

$$\mathcal{L} = -\frac{1}{4} F^A_{\mu\nu} F^{\mu\nu}_A + \sum_{\text{flavours}} \overline{\psi}_a \left(i\gamma_\mu D^\mu - m \right)_{ab} \psi_b \tag{2.1}$$

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu + g_s f^{ABC} A^B_\mu A^C_\nu \tag{2.2}$$

covariant derivative

$$\left(D_\mu \right)_{ab} = \partial_\mu \delta_{ab} - i g_s A^A_\mu t^A_{ab} \tag{2.3}$$

In order to emphasize the relevance to the measurement presented herein, the theory of Quantum Chromodynamics is discussed here from a kinematic (¡-better word?) rather than Lagrangian perspective. This is useful as jet studies help probe QCD in the soft and collinear limits. Jets are formed by the hadronization of quarks and gluons. In this thesis I present a measurement of a light quark enriched jet sample.

Consider the simplest process that could produce a quark initiated jet, a quark of energy E_q emitting a gluon of energy E_g. The probability that this will occur is a function of the gluon's energy fraction, z, and the emission angle , θ [8].

$$z = \frac{E_g}{E_q + E_g}$$

$$1 - cos\theta = \frac{m^2}{2E_q E_g}$$

Then the probability of gluon emission from the quark is :

$$P_q(z, cos\theta) dz dcos\theta = \frac{\alpha_s C_F}{\pi} \frac{dz}{z} \frac{dcos\theta}{1 - cos\theta}$$

It is useful to assume the small angle approximation, $\theta << 1$, giving:

$$P_q(z, \theta^2) dz d\theta^2 = \frac{\alpha_s C_F}{\pi} \frac{dz}{z} \frac{d\theta^2}{\theta^2}$$

Notice that the probability of emission diverges for very soft (small z) or very collinear (small θ) gluons. In the soft and collinear limits the probability can be interpretted as an expectation value for the number of soft/collinear gluons [8].

It is elucidating to rewite the probability in terms of inverse logarithms in the and intruduce the ”Lund Diagram” in order to visualize the uniform distribution of soft and collinear gluons in the $log\frac{1}{\theta^2}, log\frac{1}{z}$ space.

The primary lund plane is shown in Figure /reffig:lund [9]

$P_q(z, \theta^2)dzd\theta^2 = \frac{\alpha_s C_F}{\pi} d(log\frac{1}{z})d(log\frac{1}{\theta^2})$

Jets can also be initiated by gluons and this probability is incredibly similar :

$P_q(z, \theta^2)dzd\theta^2 = \frac{\alpha_s C_A}{\pi} d(log\frac{1}{z})d(log\frac{1}{\theta^2})$

This similarity allows us to interpret the variations in quark enriched and gluon enriched jet samples in terms of the fundamental C_F and adjoint C_A casimirs, in $SU(3)$, $C_F = \frac{4}{3}$ and $C_A = 3$.

Comparing the probability of a quark to emit a gluon and that of a gluon to to emit a gluon, we can see the ratio is simply $\frac{C_A}{C_F} = \frac{9}{4}$. This has strong experimental implications since it implies gluon jets will on average be composed of about twice as many constituent particles as quark jets.

2.2 Event Generation and reconstruction

The events used for this measurement were reconstructed from data aquired by the CMS detector in the case of the data and generated using monte carlo generators PYTHIA and HERWIG in the case of the generated data.

2.2.1 Monte Carlo Event Generators

Monte Carlo (MC) event generators are tools used by both experimental and theoretical physicists to simulate different physical processes in order to make predictions and prepare future experiments. The main tasks of such generators are to calculate matrix elements of the relevant hard processes but they must also describe parton showering, hadronization and underlying event. MC can be utilized to extrapolate data measurements beyond the acceptance of the detector or in the case of this thesis it is used in the unfolding process to correct the data for detector efficiency and resolution.

MC generators provide an ensemble of generated events which realistically describe the theoretical prediction for the physics process in question. Each individual generator implements a slighly different scheme in order to approximate the neccessary calculations for the factorization and renormalization scales relevant to a process. These variations mean that the choice of MC generator will have a slight effect on the generated distributions. For this reason, in the analysis described herein we compared results from 2 different generators: PYTHIA and HERWIG.

PYTHIA is a very commonly used general purpose event generator which uses the parton shower approach for higher order corrections to the hard scattering matrix element.

HERWIG is another commonly used event generator, incredibly similar to PYTHIA, differing mainly in hadronization and parton showering behaviors.

2.2.2 Event Reconstruction

particle flow

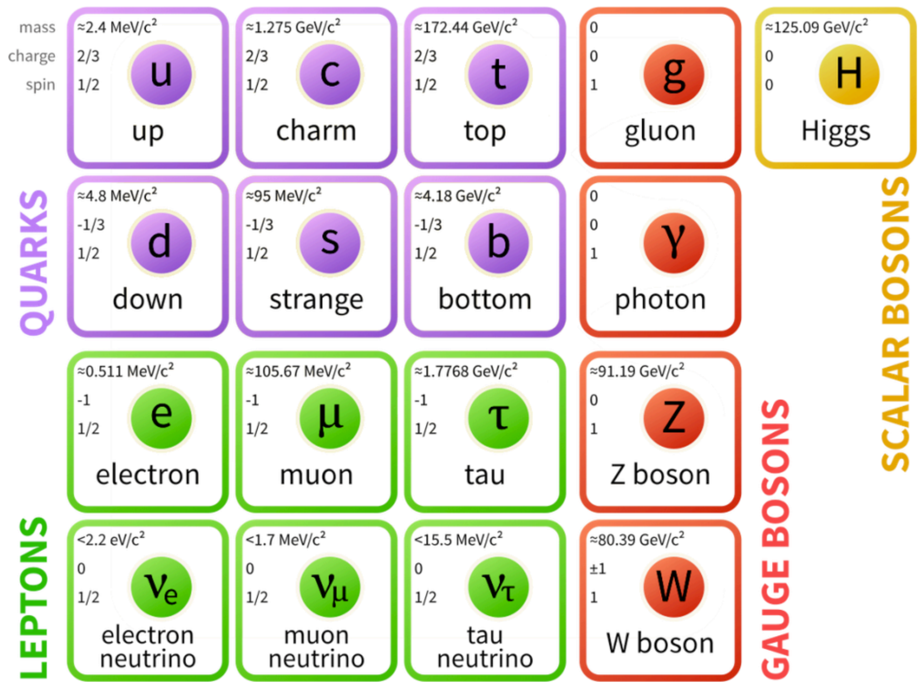

Figure 2.2: Fundamental particles of the Standard Model [7].

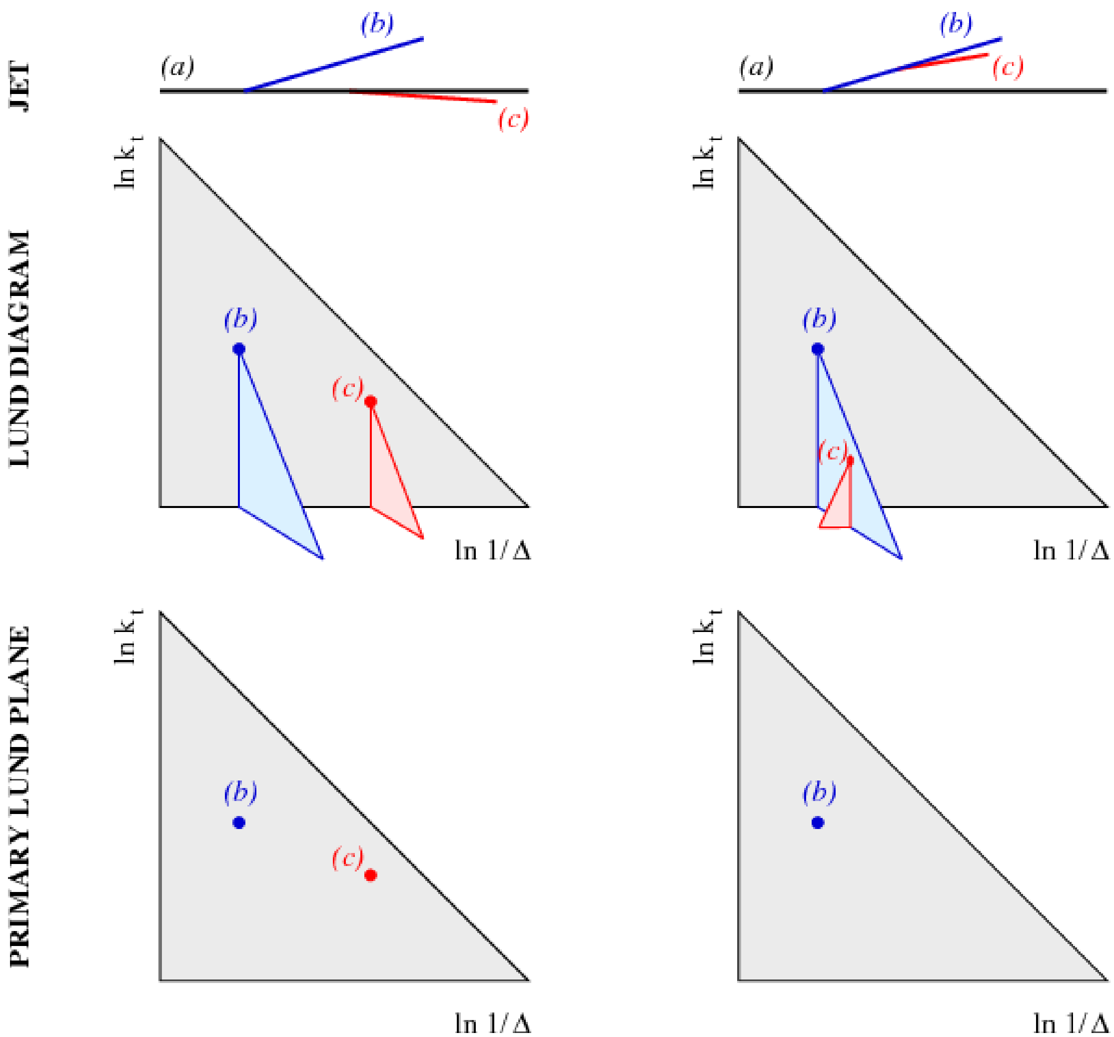

Figure 2.3: The primary and secondary lund planes for 2 example jets [9].

Chapter 3

Jets in Proton-Proton Collisions

A jet is a collimated grouping of hadrons usually associated with the production of a parton, quark or gluon, in this case initiated by the hard scatter of 2 constituent partons from protons. The initial parton radiates other quarks and gluons, called the ”Parton Shower” and all color charged particles fragment into hardons, mainly pions and kaons, before reaching the detector.

Studies of jets at LHC are complicated by experimental complexities such as ”underlying event”, other partons from the same proton iteracting and depoiting energy in the same region of the detector. ”Pileup” is also increasingly relevant, like ”underlying event” but initiated from other proton interactions from this or a previous bunch since LHC collides 10^{11} protons in bunches every 25 nanoseconds.

Lastly any measurement is limited by the resolution of the measurement device and any detector effects which are disentangled from the signal by ”unfolding” the reconstructed distributions back to generator level.

While jets are often used as simple proxies for the quark or gluon from which they originated, the structure of the radiation pattern of the hard scatter is encoded within the jet's constituent particles [10]. Jet studies are essential for a complete understanding of proton-proton interactions since the majority of interesting physics signatures contain a color charged parton in the final state. This chapter covers the basics of jet physics at LHC, from the algorithms used for clustering the constituent particles in experimental data to the language and calculations of the theory.

3.1 Jet Clustering Algorithms

Ideally, a jet represents a quark [gluon] parton however a more precise definition :

"A phase space region (as defined by an unambiguous hadronic fiducial cross section measurement) that yields an enriched sample of quarks [gluons] (as interpreted by some suitable, though fundamentally ambiguous, criterion)" [11] as defined at the Les Houches confernece in 2015 .

I will discuss one class of "suitable, though fundamentally ambiguous" criteria for defining jets, known as sequential recombination algorithms. These algorithms take pairs of particles and successively combine them into 1 particle, in a way which in intended to reconstruct the successive branchings of partons within the jet as described by perturbative QCD [12].

In sequential recombination algorithms a distance metric, d_{ij}, is defined between all particle pairs, these pairs are then sequentially combined in order of increasing distance.There exist three popular algorithms in this class, all of the following can be described by the equation [13]:

$$d_{ij} = \min\left(p_{ti}^{2p}, p_{tj}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2} \quad \Delta R_{ij}^2 = \left(y_i - y_j\right)^2 + \left(\phi_i - \phi_j\right)^2$$
$$d_{iB} = p_{ti}^{2p} \tag{3.1}$$

Depending on the value chosen for p, this equation can produce a variety of clusterings, herein I discuss the 3 popular choices $p = [1, 0, -1]$ refering to them by their names ; KT [13], Cambridge/Aachen [14] and Anti-KT algorithms [15], respectively.

KT - ¿ C/A -¿ Anti-KT [16]

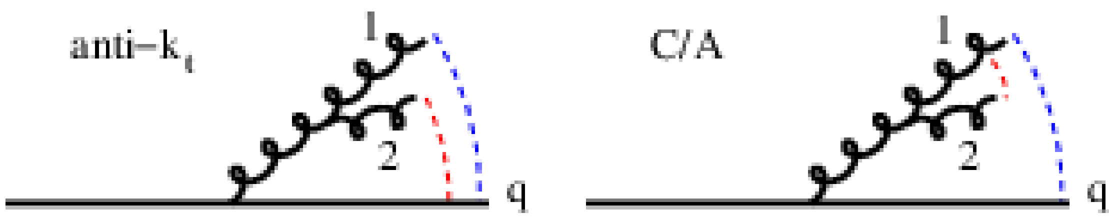

Figure 3.1: How clustering follows radiation pattern for different algorithms [9].

3.2 Jet Structure

The picture is not so simple, we have pileup and underlying event. We analyze the entire radiation pattern and from its structure can determine the probability that a gluon or quark (even the type of quark) initiated this jet there exist many methods for going about these tasks.

Quark/Gluon discrimination is discussed in section QUARK GLUON SECTION and can be done in many ways such as examining the gluon emission spectrum as visualized in the lund plane as mentioned in the introduction. In this section quark jet discrimination is discussed, specifically how the higher order corrections to the hard process are nessessary in order to properly predict measurements in the non-perturbative regime.

NLO + NLL + NP to match data in non-perturbative regime

[17]

[10]

3.3 Jet Grooming

Soft-drop

other grooomers

image of soft drop compared to others

image of soft drop grooming tree

3.4 Jet Production Cross Sections

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_n f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \frac{1}{2\hat{s}} \left|\mathcal{M}_{ab\to n}\right|^2 (\Phi_n; \mu_F, \mu_R) \quad (3.2)$$

3.5 Jet and Soft Functions in Soft-Collinear Effective Theory

Below the factorization structure of the double differential jet production cross section is displayed in the context of Soft-Collinear Effective Theory, SCET, following the framework for inclusive jet production $pp \to jet + X$ developed in for jets of

$$\frac{d\sigma}{dp_T dm} = \sum_{abc} f_a(x_a, \mu) \otimes f_b(x_b, \mu) \otimes H_{ab}^c(x_a, x_b, p_T/z, \mu) \otimes \mathcal{G}_c(z, p_T R, m, \mu, z_{\text{cut}}, \beta) \quad (3.3)$$

Image of factorization in this context

[18] discuss groomed jets in this context

comparing groomed (Jet) to ungroomed (Jet+Soft)

3.6 Jets Initited by Quarks and Gluons

earlier discussed $C_F = \frac{4}{3}$ and $C_A = 3$

CITE A Theory of Quark vs. Gluon Discrimination

Dijets make quark gluon admixture Z+Jets make mostly light quark jets, studied here and in 7 TeV analysis (1 D unfolding there and no soft drop)

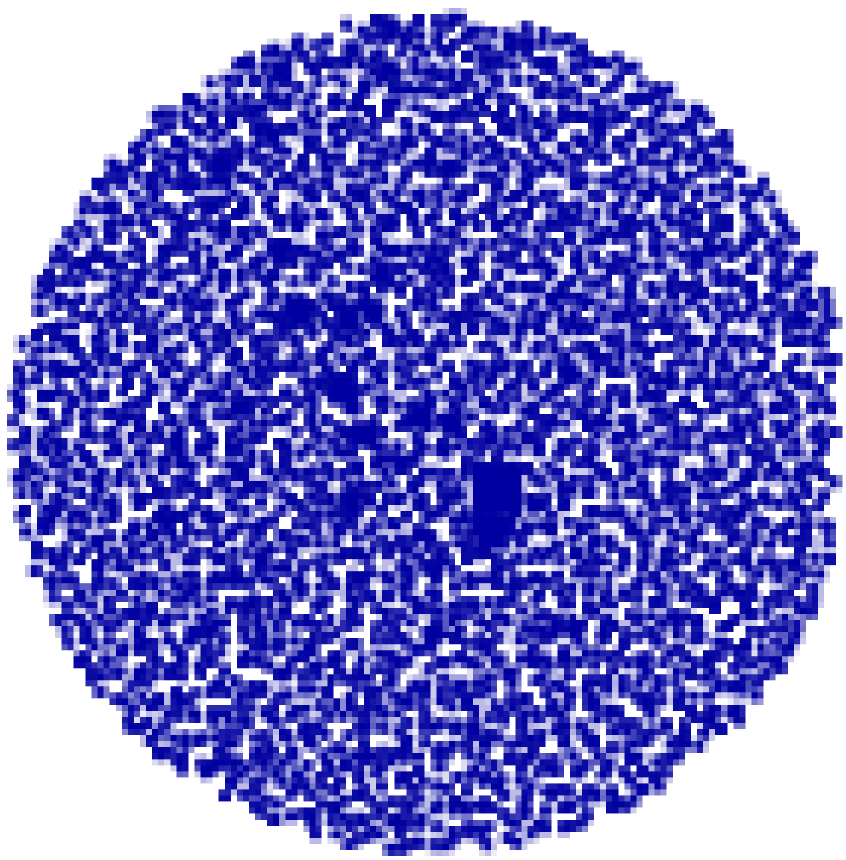

Figure 3.2: How clustering looks for Anti-Kt, circular pattern makes pileup and underlying event subtraction more simple for experimentalists [9].

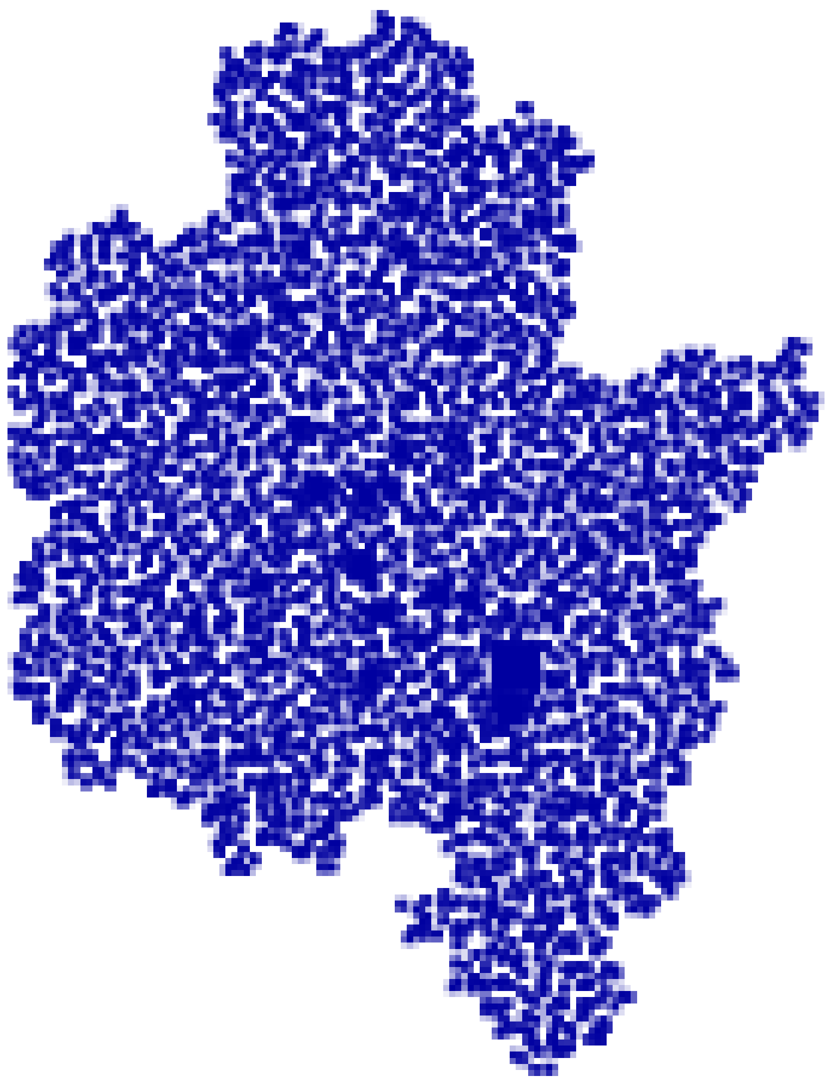

Figure 3.3: C/A, not circular, shaped like radiation pattern [9].

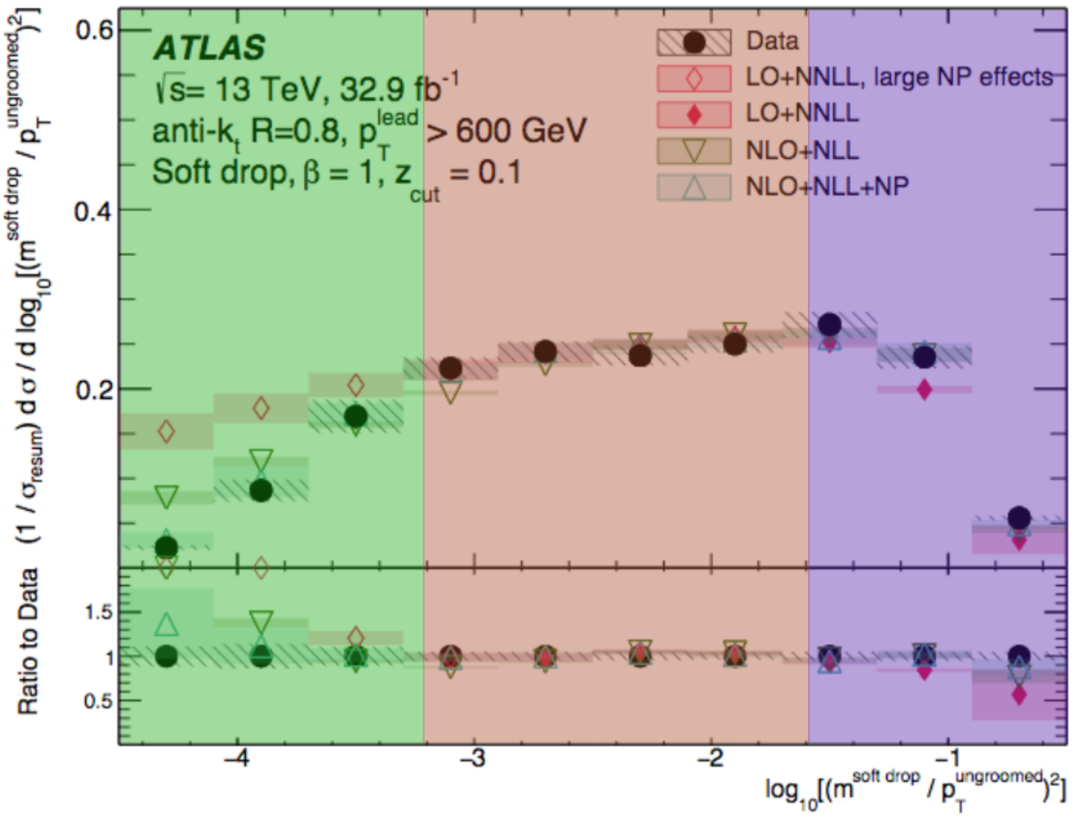

Figure 3.4: ATLS DIJets Rho result [9].

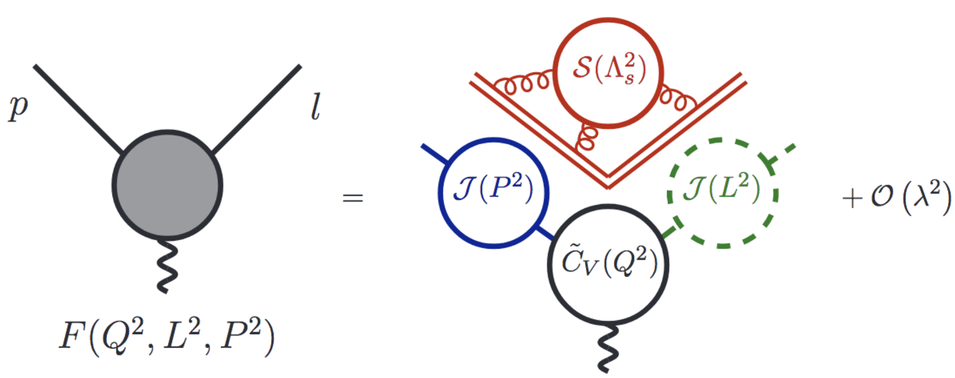

Figure 3.5: Fatorization of the energy scales in a hard scatter interaction according to SCET [18].

Chapter 4

CMS Experiment at LHC

4.1 The Large Hadron Collider

The LHC is the largest machine created my mankind to date and currently the world's highest energy particle accelerator, it accelerates and collides bunches of 10^{11} protons at a time which collide at combined center-of-mass energy of 13 TeV.

4.2 The CMS Detector

The Compact Muon Solenoid (CMS) detector was used to collect the data presented in this thesis, it is one of two large general purpose detectors at the LHC. CMS experiment has recorded 162 fb^{-1} integrated luminosity in the dataset presented in this thesis, collected during Run 2 of LHC.

The Compact Muon Solenoid, CMS, is one of 4 detectors that measure collisions of protons and lead ions produced by the Large Hadron Collider, LHC, at CERN. CMS is the smaller of the 2 large general-purpose detectors, the other being ATLAS. The most notable feature of the detector is it's powerful 3.8 Tesla solenoid magnet, the largest superconducting magnet ever built, as of the year 2011.

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, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters, made of steel and quartz-fibres, extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

Chapter 5

Measurement of the differential jet production cross section with respect to jet mass and transverse momentum in Z + Jet events from pp collisions at $\sqrt{s} = 13$ TeV

Abstract.

We present a measurement of the differential jet production cross section as a function of the jet mass and transverse momentum in events with a Z + Jet topology, with and without a jet grooming algorithm applied. For ungroomed jets, leading-order and next-to-leading order QCD Monte Carlo programs are found to predict the jet mass spectrum in the data reasonably well, with some disagreement at very low and very high masses. For groomed jets, the agreement between the Monte Carlo programs and the data improves overall, and

extends lower in jet mass due to the removal of soft and collinear portions of the jet. First-principles theoretical calculations of the groomed jet mass are in preparation by theorists to be compared to the data.

5.1 Introduction

We present a measurement of the differential production cross section of Z+jets events as a function of the jet mass and transverse momentum (pt). The cross section is presented for events before and after the jets are groomed with the "soft drop" procedure [19], using multiple values of the tunable parameters β and z_{cut}. Softdrop iteratively declusters a jet j with distance parameter R into two subjets, j_1 and j_2. If the softdrop condition

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \cdot (\frac{\Delta R_{12}}{R})^{\beta} \tag{5.1}$$

is met, then the procedure stops and j is the final jet. Otherwise, the declustering continues - the higher pt subjet is relabeled as j and the lower pt one is dropped. By design, this condition fails for wide-angle soft radiation, which is therefore removed by the soft drop procedure. The tunable parameters, β and z_{cut}, control the degree of jet grooming: β tunes the algorithm's sensitivity to wide-angle radiation, while z_{cut} sets the energy scale of the grooming. In the case of $\beta \to \infty$, an ungroomed jet is returned. In the $\beta = 0$ case, the soft drop procedure is identical to the "modified mass drop tagger" (MMDT) from Ref. [20]. The soft drop algorithm removes soft and wide-angle radiation from jets in a very theoretically controlled manner, making it suitable to separate the "hard" and "soft" parts of the jet. Specifically, the soft drop algorithm can remove non-global logarithms from correlations of radiation within and between jets, which are extremely difficult to compute theoretically [21, 20, 19, 22, 23, 24].

Comparing the production cross section for groomed and ungroomed jets separately allows us to gain sensitivity to both the “hard” and “soft” jet physics. The groomed cross section can be directly compared to theoretical calculations of the jet mass now and in the future, which is a very active area of theoretical research at this time [25, 26, 27, 28, 29, 30, 31, 32, 33]. Furthermore, separating the hard and soft jet physics allows a deeper understanding of the various effects involved in QCD radiation. In particular, Ref. [32] calculates the groomed jet mass at next-to-next-to-leading order using soft colinear effective theory, matched to a parton shower at leading order using `MCFM` [34, 35], and the authors of Ref. [20] have provided a next-to-leading logarithm calculation with traditional perturbative QCD, matched to a parton shower at leading order, also using `MCFM`. We compare these theoretical predictions to our data in this paper for the first time in this channel at CMS. Both CMS and ATLAS have similar measurements in a dijet sample at Ref. [36, 37].

The analysis strategy is similar to that of Ref. [36]. However, there are several differences. As in that paper, the cross section is now unfolded in both jet mass and pt. However, while the previous measurement considered only one value for the soft drop parameter β, this analysis considers several. We apply the soft drop algorithm to compare directly to theoretical computations. Additionally, we not only measure the cross section as a function of mass, but also as a function of dimensionless mass, $\rho = 2log(m/(ptR))$, as is also done in the previously mentioned ATLAS measurement. The dimensionless mass ρ only weakly depends on pt, unlike mass, which is highly correlated. Additionally, the use of this variable aids in the separation of fixed order, perturbative and non-perturbative effects. Finally, we also present the normalized differential cross section. We compute the cross sections normalized per p_{T} bin (the "normalized" cross section) with respect to the jet pt and jet mass by unfolding a binned two-dimensional distribution in pt and mass with widths Δpt and Δm, respectively.

The normalized differential cross section

$$\frac{1}{d\sigma/dpt}\frac{d^2\sigma}{dpt\,dm} = \frac{1}{N/\Delta pt}R(\frac{N_{ij}}{\Delta pt\,\Delta m}) \tag{5.2}$$

where N is the total number of Z+jets events in our selection, N_{ij} is the number of such events in pt bin i and mass bin j, and $R(\alpha)$ is the unfolding procedure applied to the two-dimensional distribution α.

5.2 Data and MC

The data used are from proton-proton collisions in 2016, 2017, and 2018 with 25 ns bunch spacing. The datasets are summarized in Table **??**.

The approved luminosity sections used for 2016, 2017, 2018 Runs A – C, and 2018 Run D, respectively, are listed below:

- `Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt`
- `Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON.txt`
- `Cert_314472-325175_13TeV_PromptReco_Collisions18_JSON.txt`
- `Cert_314472-325175_13TeV_PromptReco_Collisions18_JSON.txt`

In this analysis, Drell-Yan Monte Carlo (MC) events were simulated with MadGraph [38] and showered with two different generators.

Data sample	Cross section (pb^{-1})
2016	
/SingleMuon/Run2016*-17Jul2018-v1/MINIAOD	35.9
/SingleElectron/Run2016*-17Jul2018-v1/MINIAOD	35.9
2017	
/SingleMuon/Run2017*-31Mar2018-v1/MINIAOD	41.5
/SingleElectron/Run2017*-31Mar2018-v1/MINIAOD	41.5
2018 Runs A – C	
/SingleMuon/Run2018*-17Sep2018-v1/MINIAOD	27.9
/EGamma/Run2018*-17Sep2018-v1/MINIAOD	27.9
2018 Run D	
/SingleMuon/Run2018D-PromptReco-v2/MINIAOD	31.3
/EGamma/Run2018D-PromptReco-v2/MINIAOD	31.1

Table 5.1: Data samples used in the analysis.

The first generator is PYTHIA 8 [**?**], for which inclusive as well as ($\hat{p}_T$) binned samples were used.

An ensemble of samples, each varying the hard-scatter transverse momentum ($\hat{p}_T$) from 70 GeV to Infinity, was used to fill the response matrices for the measurement. The samples are :

- 2016: `/DYJetsToLL_M-50_HT-*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/`
- 2017 and 2018: `/DYJetsToLL_M-50_HT-*_TuneCP5_13TeV-madgraphMLM-pythia8/`

The underlying event tune is `CUETP8M1` [39] for 2016 and `CP5` [40] for 2017 and 2018. These samples were generated in the

- `RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3`
- `RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14`
- `RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15`

MC campaigns for 2016, 2017, and 2018, respectively. A second inclusive PYTHIA8 sample was used to evaluate the physics model uncertainty, for which the underlying event tune is `CUETP8M1`. It generated with `CMSSW 9.4.X` in the `RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3` MC campaign.

The second generator is HERWIG++ [41] with tune `CUETP8M1`, which is also used to generate a single sample which was also used to evaluate the physics model uncertainty. This sample was generated with `CMSSW 9.4.X` in the `RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3` MC campaign [This was used because there was only a 2016 version available].

Table 5.2 summarizes the Drell-Yan samples, the number of events, and the cross sections.

The backgrounds considered in this analysis comprise $t\bar{t} + jets$ and diboson (WW, WZ, and ZZ) samples. The $t\bar{t} + jets$ samples were simulated with MADGRAPH and showered with PYTHIA 8, while the diboson samples with simulated and showered with PYTHIA 8. The tunes and MC campaigns used were the same as for the Drell-Yan samples. Table 5.3 summarizes the background samples, the number of events, and the cross sections.

The reconstruction used for the data and MC was `CMSSW 9.4.X` (`CMSSW 10.0.X`) for 2016 and 2017 (2018), with private extended `NANOAOD` based on central `NANOAODv4`. This corresponds to the following `MINIAOD` versions:

- 2016: v3
- 2017: v2
- 2018: v1

MC sample	Number of Events	Cross section (pb), LO
2016		
HERWIG++	29883521	1091.1
PYTHIA 8 Tune CUETP8M1	96531428	14420
PYTHIA 8 Tune CUETP8M1 70-100	9691660	175.3
PYTHIA 8 Tune CUETP8M1 100-200	2751187	147.4
PYTHIA 8 Tune CUETP8M1 100-200 extension	8265899	147.4
PYTHIA 8 Tune CUETP8M1 200-400	962195	40.99
PYTHIA 8 Tune CUETP8M1 200-400 extension	8646942	40.99
PYTHIA 8 Tune CUETP8M1 400-600	1070454	5.678
PYTHIA 8 Tune CUETP8M1 400-600 extension	8655207	5.678
PYTHIA 8 Tune CUETP8M1 600-800	8292957	1.363
PYTHIA 8 Tune CUETP8M1 800-1200	2673066	0.6759
PYTHIA 8 Tune CUETP8M1 1200-2500	596079	0.116
PYTHIA 8 Tune CUETP8M1 2500-Inf	399492	0.002592
2017		
PYTHIA 8 Tune CP5 70-100	9344037	143
PYTHIA 8 Tune CP5 100-200	10235418	164.4
PYTHIA 8 Tune CP5 100-200 extension	3950339	164.4
PYTHIA 8 Tune CP5 200-400	10728447	49.8
PYTHIA 8 Tune CP5 200-400 extension	1200863	49.8
PYTHIA 8 Tune CP5 400-600	9533635	6.8
PYTHIA 8 Tune CP5 400-600 extension	1124294	6.8
PYTHIA 8 Tune CP5 600-800	8153358	1.8
PYTHIA 8 Tune CP5 800-1200	3089861	0.85
PYTHIA 8 Tune CP5 1200-2500	625517	0.018
PYTHIA 8 Tune CP5 2500-Inf	404986	0.0036
2018		
PYTHIA 8 Tune CP5 70-100	10019684	143
PYTHIA 8 Tune CP5 100-200	11530510	164.4
PYTHIA 8 Tune CP5 200-400	11210867	49.8
PYTHIA 8 Tune CP5 400-600	9697098	6.8
PYTHIA 8 Tune CP5 400-600 extension	9358053	6.8
PYTHIA 8 Tune CP5 600-800	8862104	1.8
PYTHIA 8 Tune CP5 800-1200	3138129	0.85
PYTHIA 8 Tune CP5 1200-2500	536416	0.018
PYTHIA 8 Tune CP5 2500-Inf	427051	0.0036

Table 5.2: List of Monte Carlo samples used. The number of generated events and the total cross section are also provided for each subsample.

MC sample	Number of Events	Cross section (pb)
2016		
TTJets PYTHIA 8 Tune CUETP8M1	10199051	831.76, NNLO
WW PYTHIA 8 Tune CUETP8M1	6988168	118.7, NNLO
WZ PYTHIA 8 Tune CUETP8M1	2997571	47.13, NLO
ZZ PYTHIA 8 Tune CUETP8M	998034	16.523, NLO
2017		
TTJets PYTHIA 8 Tune CUETP8M1	8026103	831.76, NNLO
WW PYTHIA 8 Tune CUETP8M1	7765828	118.7, NNLO
WZ PYTHIA 8 Tune CUETP8M1	3928630	47.13, NLO
ZZ PYTHIA 8 Tune CUETP8M	1925931	16.523, NLO
2018		
TTJets PYTHIA 8 Tune CUETP8M1	10244307	831.76, NNLO
WW PYTHIA 8 Tune CUETP8M1	7850000	118.7, NNLO
WZ PYTHIA 8 Tune CUETP8M1	3885000	47.13, NLO
ZZ PYTHIA 8 Tune CUETP8M	1979000	16.523, NLO

Table 5.3: List of Monte Carlo samples used. The number of generated events and the total cross section are also provided for each subsample.

5.3 Trigger

The data are collected with single-lepton triggers, with muons (electrons) of $pt > 29(37)$ GeV using the $IsoMu27$ ($Ele35_WPTight_Gsf$ || $Photon200$) triggers, as recommended by the muon (egamma) POG. All triggers used for this analysis are prescaled to 1.

5.4 Reconstruction and Selection

The event reconstruction is based on the CMS Particle Flow (PF) algorithm [42, 43, 44], which takes into account information from all subdetectors, including charged particle tracks from the tracking system, energy deposits in the electromagnetic and hadronic calorimeters and tracks reconstructed in the muon chambers. Given this information, all particles in the event are reconstructed as electrons, muons, photons, charged hadrons or neutral hadrons. Charged hadrons associated with pileup vertices are removed from consideration, referred to as "charged hadron subtraction" (CHS).

Hadronic jets are clustered from particle flow inputs using the anti-k_T algorithm [15] using the FASTJET 3.0 software package [45] with $R = 0.8$ (AK8 jets). Corrections based on the jet area [46] are applied to the jets to remove the energy contribution of neutral hadrons arising from pileup collisions. Further corrections are used to account for the nonlinear calorimetric response, as a function of η and p_T [**?**], derived from simulation and data-to-simulation correction factors. `L1FastJet`, `L2`, `L3`, and `L2L3Residual` AK8PFchs corrections from version `Fall15_25nsV2` were used. For the jet energy, all corrections are applied. For the jet mass, only `L2L3+Residual` corrections are applied, and the `L1` correction is not applied. This is to ensure that theoretical calculations can be directly compared. It is also the recommendation from JMAR.

The constituents of the AK8 jets are reclustered using the Cambridge-Aachen algorithm [**? ?**]. The jet grooming algorithm used is the modified mass drop tagger (MMDT) algorithm [20], also known as the "soft drop" algorithm with angular exponents $\beta = 0, 1, -1$, soft thresholds $z_{cut} < 0.05, 0.1, 0.15$, all with characteristic radius $R_0 = 0.8$ [19], giving a total of 9 different groomings. This algorithm is also used to identify two subjets within the AK8 jet.

The same reconstruction criteria are also applied to stable particles at the generator level (`GenJets`), removing neutrinos.

All reconstructed AK8 jets are required to have $p_\mathrm{T} > 200$ GeV and rapidity $|y| < 2.4$, as well as to satisfy "tight" jet identification requirements to remove detector noise: CHECK 2017 TIGHT JET reqirements

- Neutral hadron fraction < 0.90,
- Neutral EM fraction < 0.90,
- Charged hadron fraction > 0.00,
- Charged EM fraction < 0.99,
- At least two constituents,
- At least one charged hadron.

To account for the worse jet energy resolution in the data than the MC, the MC are smeared in the nominal case by the amounts listed in Ref. [**?**]. The typical values are around 5-8%. These are applied to the entire jet four-vector, including p_{T} and mass. Currently the `Fall17 V3` JER files are being used.

The jet mass can also have a different scale and resolution in data and MC. These are determined by fitting the W mass in semileptonic $\mathrm{t\bar{t}}$ events. The ratio of the fitted means (data/MC) is the jet mass scale, JMS, and the ratio of the fitted widths (data/MC) is the jet mass resolution, JMR. Traditionally, this procedure is performed on events where the W decay products merge forming a since AK8 Jet. At high momentum, above 400 GeV, this method breaks down since the b quark decay products merge into the same AK8 as the W. Using the highest mass soft-drop subjet as the W candidate, we are able to more accurately determine JMR and JMS for higher transverse momentum jets. These studies are ongoing, currently, we are using dummy values from the nanoAOD-tools modules JMR [0.1, 0.2, 0.0] and JMS [1. , 0.99, 1.01].

The jet mass resolution for ungroomed jets varies from XXX% depending on p_{T} and mass. The ungroomed jet mass resolution below XX GeV is very non-Gaussian, thus poorly estimated (see Figures **??**-**??**). FIX THIS SENTENCE For this reason, the jet mass measurement is deemed to be invalid for ungroomed masses below 20 GeV for $p_{\mathrm{T}} < 760$ GeV and below 40 GeV for $p_{\mathrm{T}} > 760$ GeV. This is not very problematic because there are very few events for ungroomed masses in this regime.

The groomed jets, on the other hand, can access jet masses very close to zero. The jet mass scale in MC is between XXX 0.8-1.0 for the groomed jets for all jet p_{T} and masses. Small deviations from unity are accounted for in the response matrix. However, the jet mass resolution for very closely-separated jet constituents becomes extremely poor. This is because of the finite spatial resolution of the detector. Since the opening angle between the constituents scales like m/p_{T}, there will be minimum masses at a given p_{T}, below which the measurement is invalid. The criterion chosen here is to require that the uncertainty of the fits to the jet mass scale be lower than 50%. This is shown as the uncertainty bands in Figs **??**-**??** and then again as the central values in Figs. **??**-**??**. With this criterion in mind, the measurement of the groomed masses are restricted to $m > 10$ GeV.

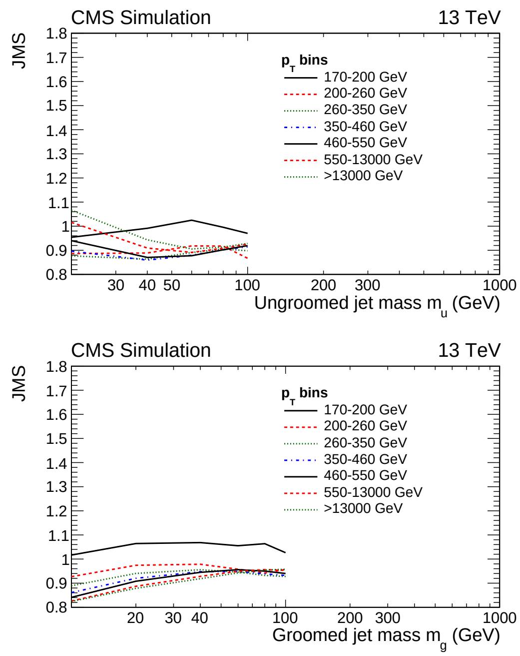

Figure 5.1: Jet mass scale in MC (mean of fits to m_{reco}/m_{gen} as a function of $p_{T_{gen}}$) for ungroomed (top) and groomed (bottom) jets in different p_T bins as a function of mass.

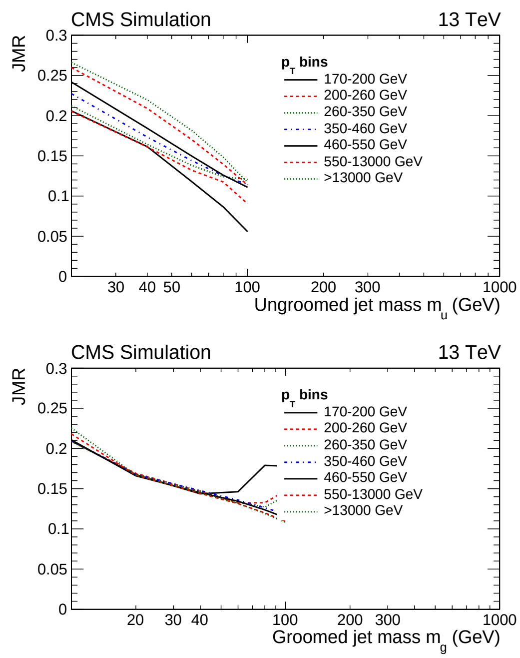

Figure 5.2: Jet mass resolution in MC (widths of fits to m_{reco}/m_{gen} as a function of $p_{\mathrm{T}gen}$) for ungroomed (top) and groomed (bottom) jets in different p_{T} bins as a function of mass.

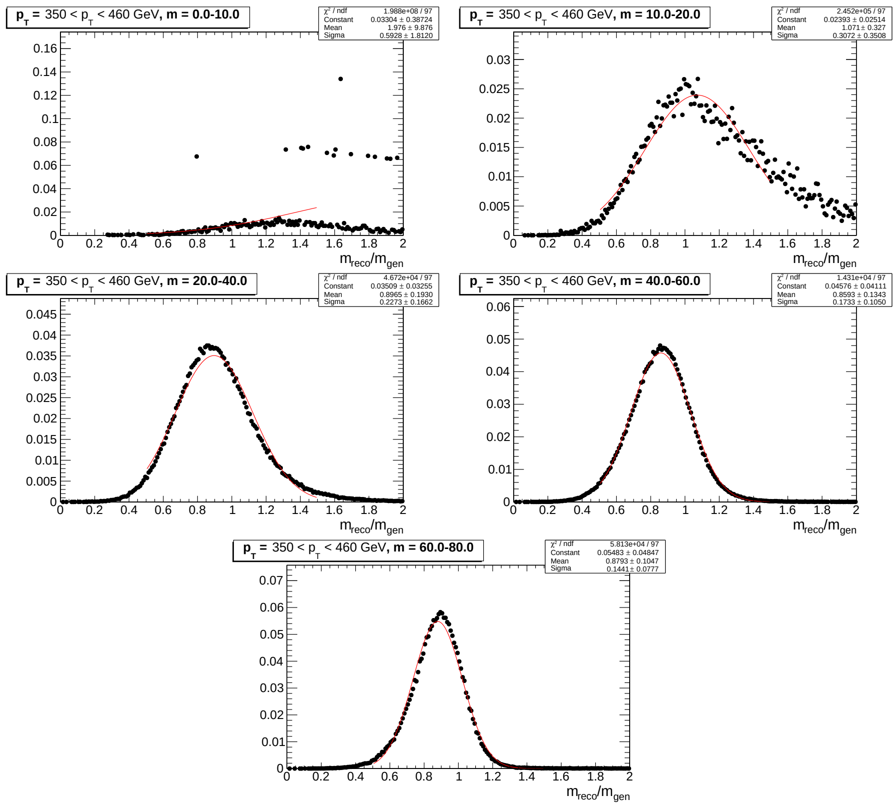

Figure 5.3: Fits to m_{reco}/m_{gen} for different m_{gen} bins are shown for ungroomed jets of p_{T} 350-460 GeV . The full fits are in Appendix **??** for separate p_{T}, m_{gen} bins.

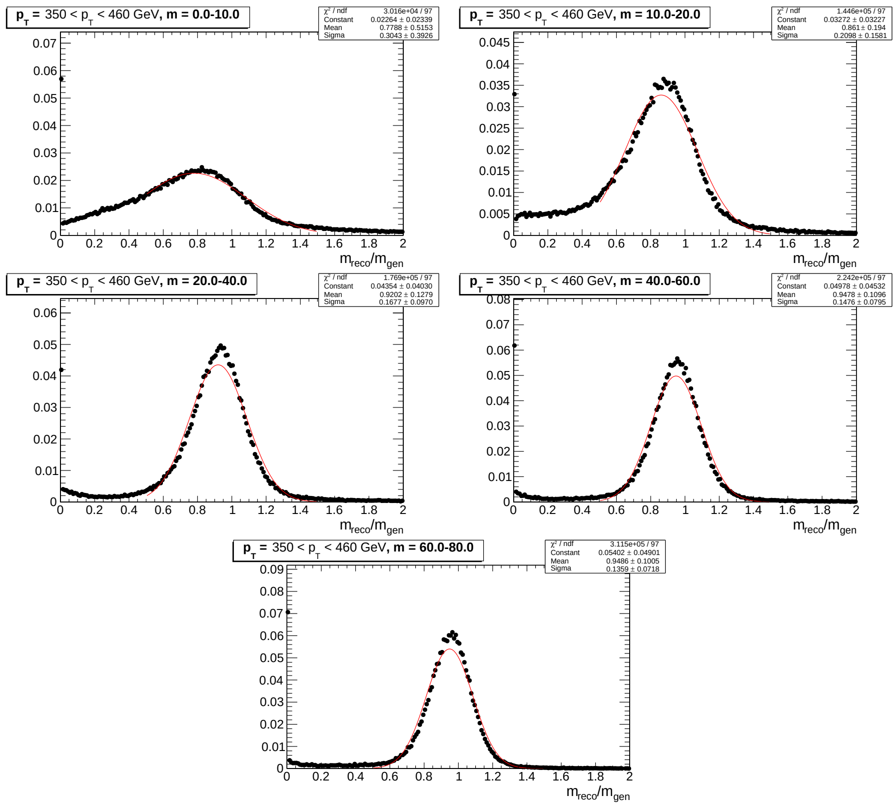

Figure 5.4: Fits to m_{reco}/m_{gen} for different m_{gen} bins are shown for groomed jets of p_{T} 350-460 GeV . The full fits are in Appendix **??** for separate p_{T}, m_{gen} bins.

5.5 Data to MC Comparisons

The jet p_{T} of the leading jet, this is the only jet considered in the event, after luminosity scaling is shown for muon events which passed any of the 3 triggers in Fig. 5.5.

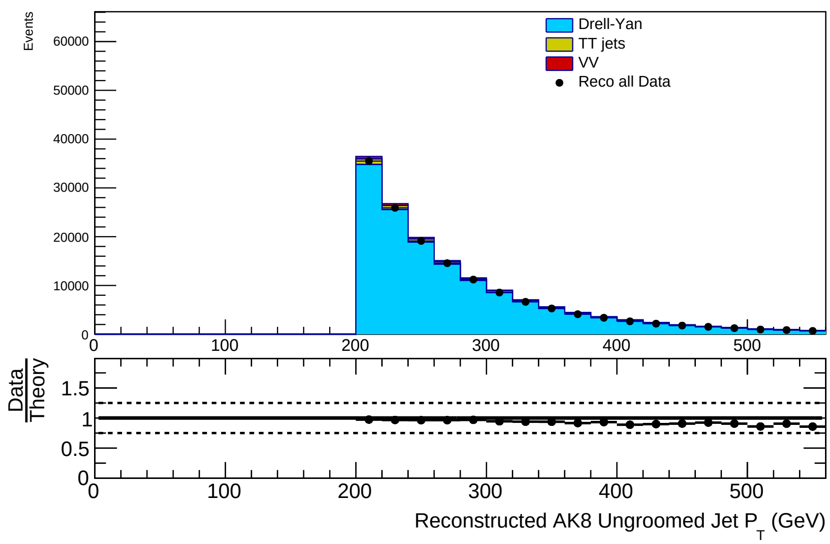

Figure 5.5: Jet p_{T} of leading jet of all data compared with PYTHIA 8 monte carlo simulated events, with luminosity scaling for 2017 data applied to the simulation.

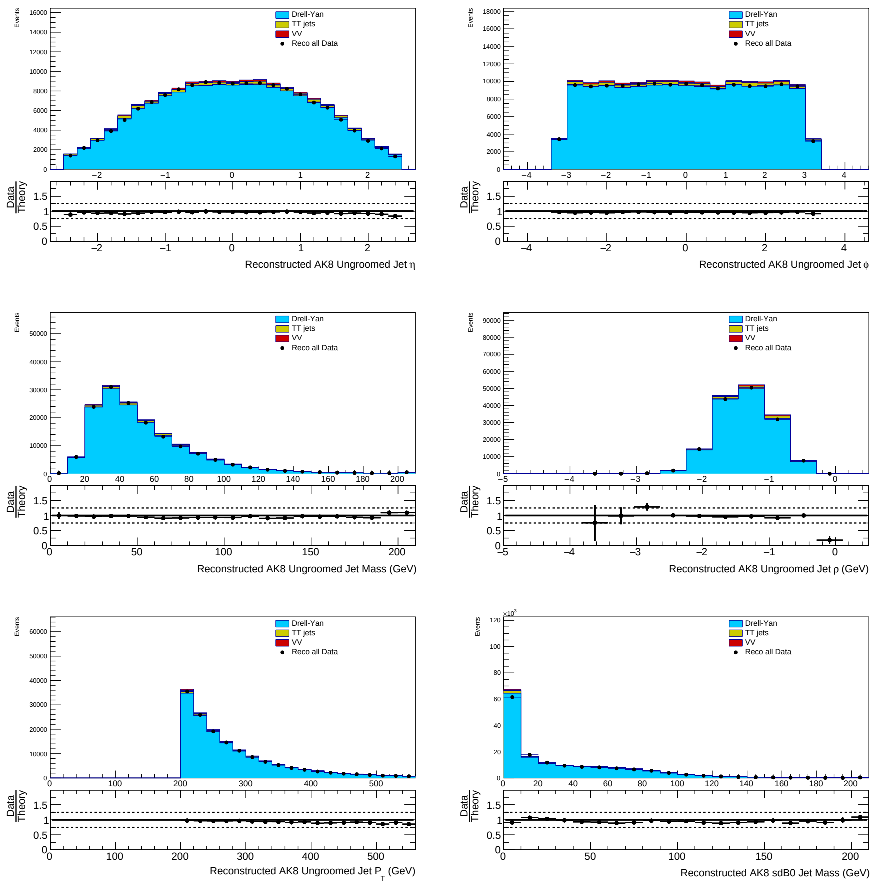

Figure 5.6: Jet pseudorapidity, azimuthal angle, mass, ρ [perhaps groomed ρ is better or we should show both side by side.] , p_{T}, and soft-drop mass (shown here for the selected case $\beta = 0$ and $Z_{cut} = 0.1$ with plots of the other groomed masses in the appendix "Data and MC comparisons for soft-dropped jet masses") for data and PYTHIA8 MC, inclusive in jet p_{T}.

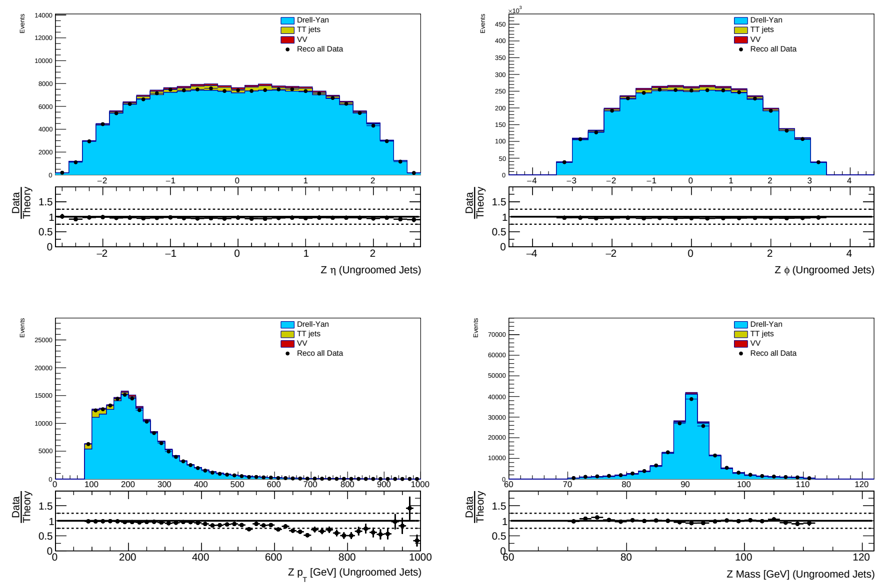

Figure 5.7: Z candidate pseudorapidity, azimuthal angle, mass, p_{T}, inclusive in jet p_{T}.

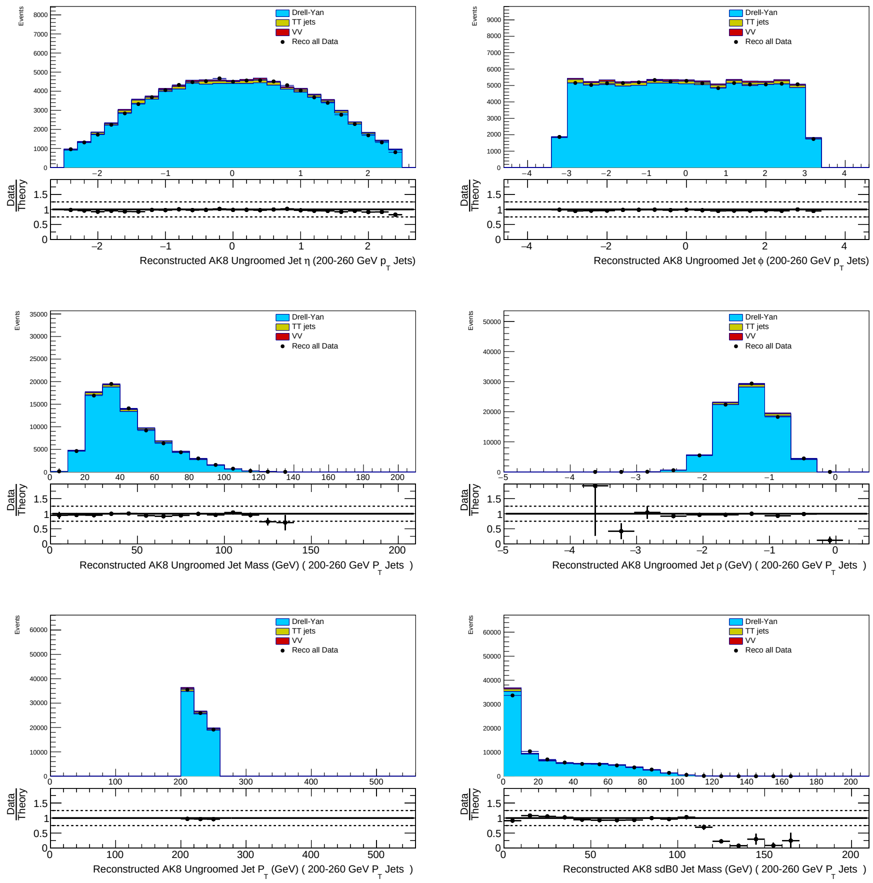

Figure 5.8: Jet pseudorapidity, azimuthal angle, mass, ρ, p_{T}, and soft-drop mass (shown here for the selected case $\beta = 0$ and $Z_{cut} = 0.1$ with plots of the other groomed masses in the appendix "Data and MC comparisons for soft-dropped jet masses") for data and PYTHIA8 MC for $p_T = 200 - 260$ GeV.

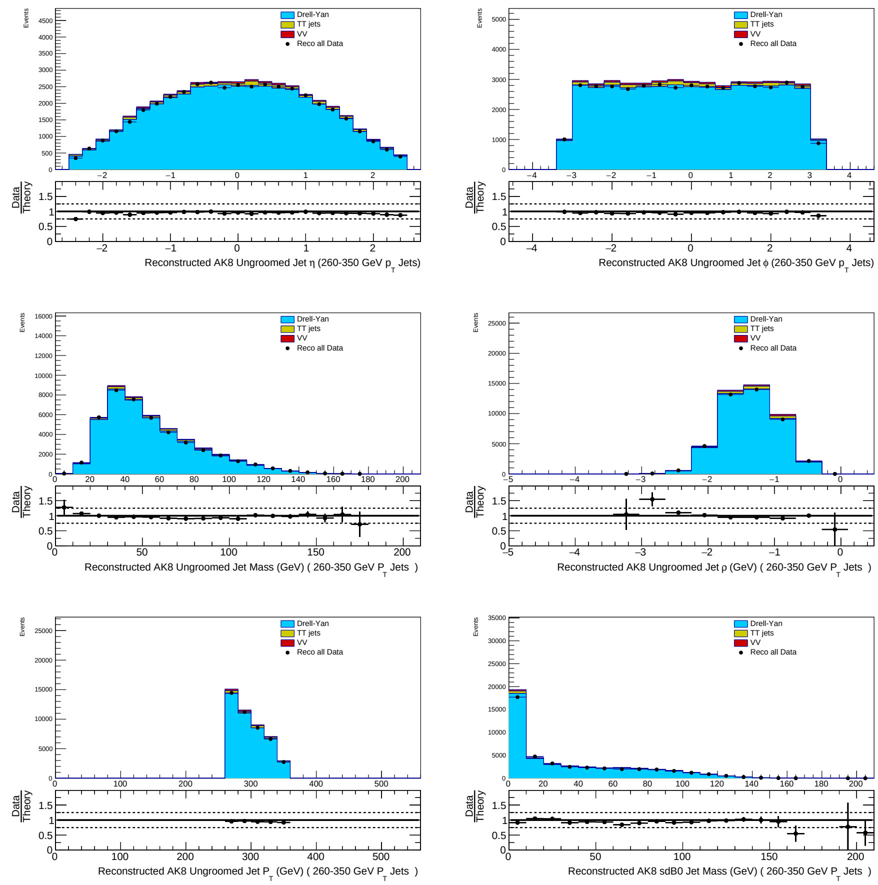

Figure 5.9: Jet pseudorapidity, azimuthal angle, mass, ρ, p_{T}, and soft-drop mass (shown here for the selected case $\beta = 0$ and $Z_{cut} = 0.1$ with plots of the other groomed masses in the appendix "Data and MC comparisons for soft-dropped jet masses") for data and PYTHIA8 MC for $p_T = 260 - 350$ GeV.

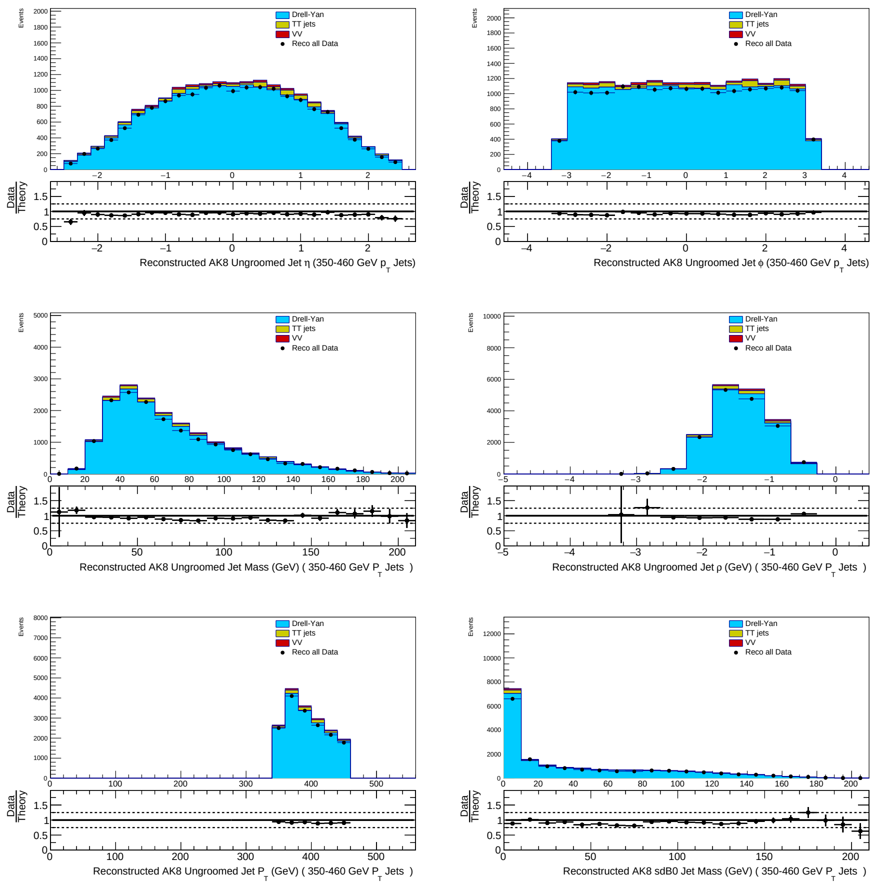

Figure 5.10: Jet pseudorapidity, azimuthal angle, mass, ρ, p_{T}, and soft-drop mass (shown here for the selected case $\beta = 0$ and $Z_{cut} = 0.1$ with plots of the other groomed masses in the appendix "Data and MC comparisons for soft-dropped jet masses") for data and PYTHIA8 MC for $p_T = 350 - 460$ GeV.

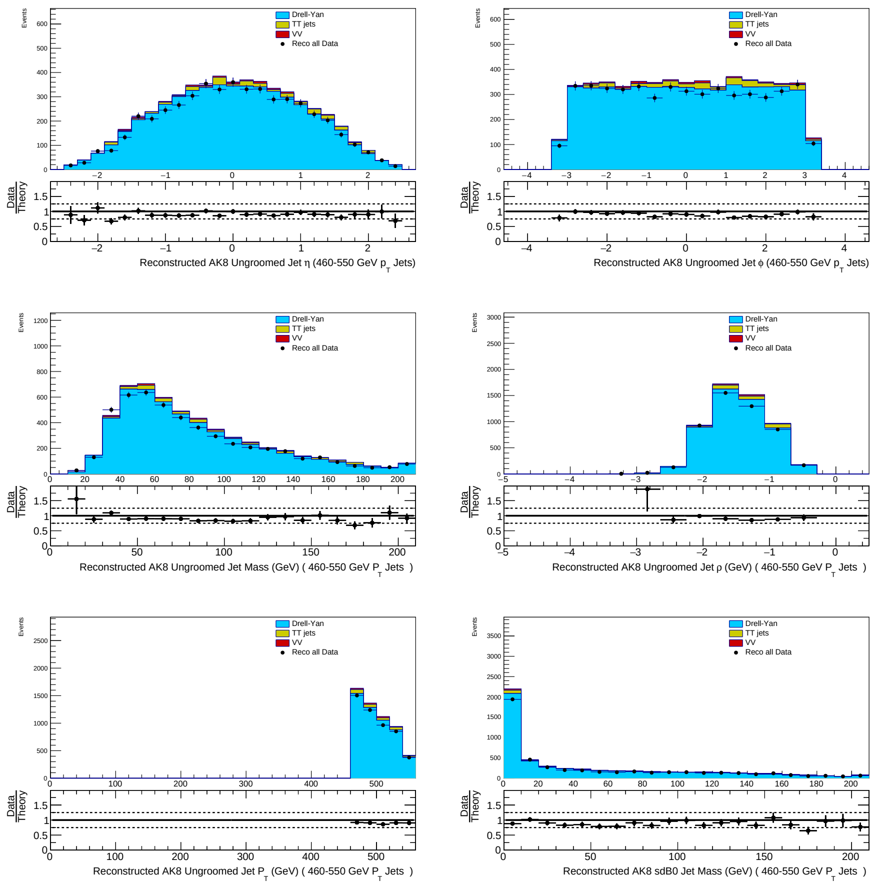

Figure 5.11: Jet pseudorapidity, azimuthal angle, mass, ρ, p_{T}, and soft-drop mass (shown here for the selected case $\beta = 0$ and $Z_{cut} = 0.1$ with plots of the other groomed masses in the appendix "Data and MC comparisons for soft-dropped jet masses") for data and PYTHIA8 MC for $p_T = 460 - 550$ GeV.

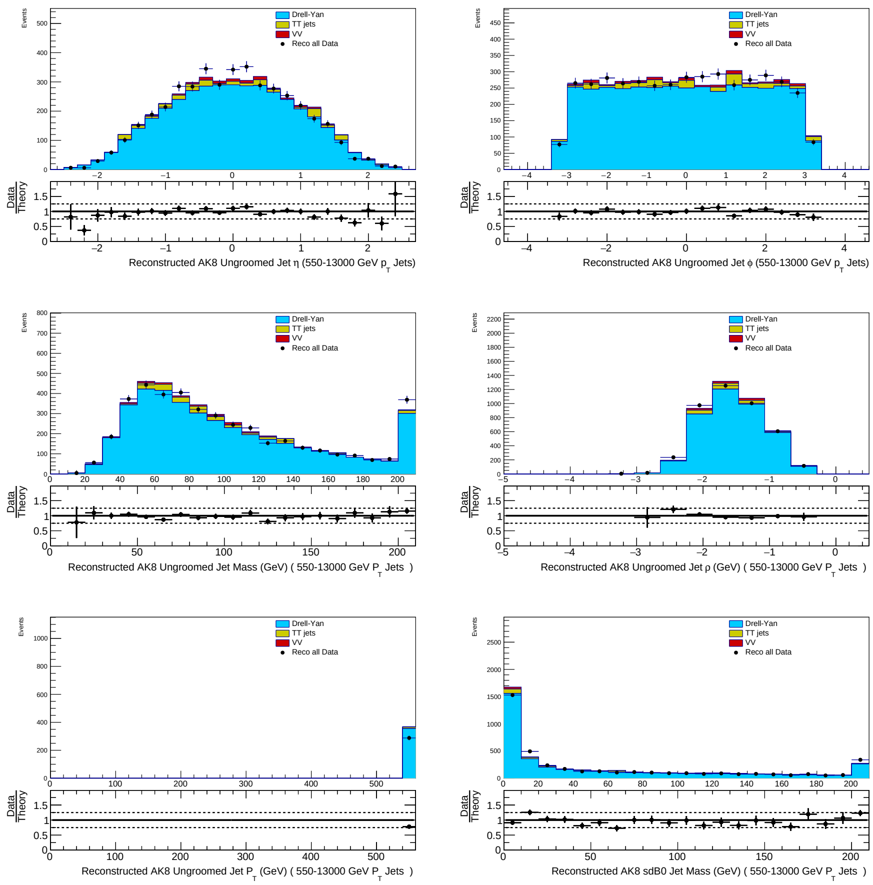

Figure 5.12: Jet pseudorapidity, azimuthal angle, mass, ρ, p_{T}, and soft-drop mass (shown here for the selected case $\beta = 0$ and $Z_{cut} = 0.1$ with plots of the other groomed masses in the appendix "Data and MC comparisons for soft-dropped jet masses") for data and PYTHIA8 MC for $p_T = 550 - 1300$ GeV.

5.6 Detector Response

The two-dimensional detector response is generated from MC simulation in bins of jet p_{T} and mass (Later we will include these distribtions for jet p_{T} and ρ) . For the groomed jet measurement, the ungroomed jet p_{T} and the groomed jet mass are used (here results for $\beta = 0$ and $z_{cut} = 0.1$ are shown while alternatively groomed distributions are given in the appendix). This is because the theoretical predictions we will compare to use the ungroomed jet p_{T}, which is a Sudakov safe observable unlike the soft-drop groomed pt. The central value is constructed using PYTHIA8 with all systematic variations at their nominal values. The systematic variations are treated by unfolding the data or MC with new responses with the systematic variations applied, then taking the full difference from the nominal value. The response matrix uncertainties are symmetrized. The matrix is visualized by normalizing each row to unity.

The p_{T} bins are $200, 260, 350, 460, 550, 13000$ GeV, they were chosen to match those of the previous measurement in the dijet channel in order to enable later comparisons of jets with different flavor compositions, here we present a like-quark enriched sample. We did not have sufficient statistics at high Pt to maintain the same binning as the dijet case above 550 GeV jet Pt. The response matrix is shown for ungroomed jets and groomed jets in Figures **??** and **??**, respectively. The jet mass bins are chosen as a trade-off between the approximate detector resolution of $\sim$10 GeV and the purity and stability. To avoid many bins below 10%, the binning chosen is initially finely binned and then becomes broader.

The purity is defined as the fraction of events in a reconstructed bin that are generated from the same bin. (That is, the diagonal value divided by the row it is in). The stability is defined as the fraction of events in a generated bin that are reconstructed in the same bin. (That is, the diagonal value divided by the column it is in). To avoid unphysical regions (where the jet mass is above what can occur given a jet p_{T}), the purity and stability are restricted to the regions where $m < p_{\mathrm{T}} * R/\sqrt{2}$. These are shown in Figures 5.17and Figures **??** for ungroomed and groomed jets, respectively.

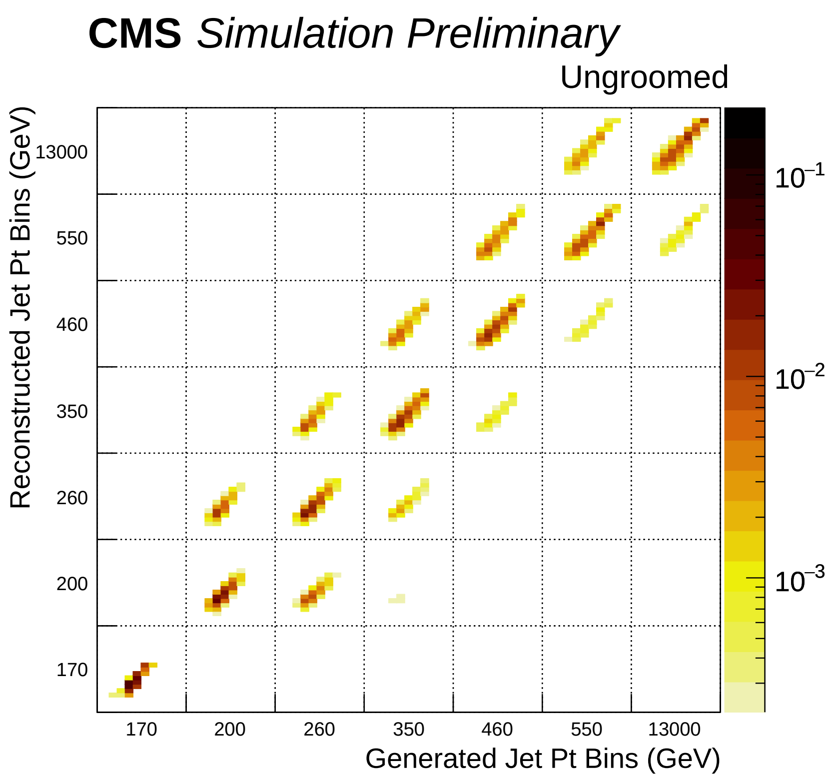

Figure 5.13: Two-dimensional response matrix for ungroomed jets. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV). While the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

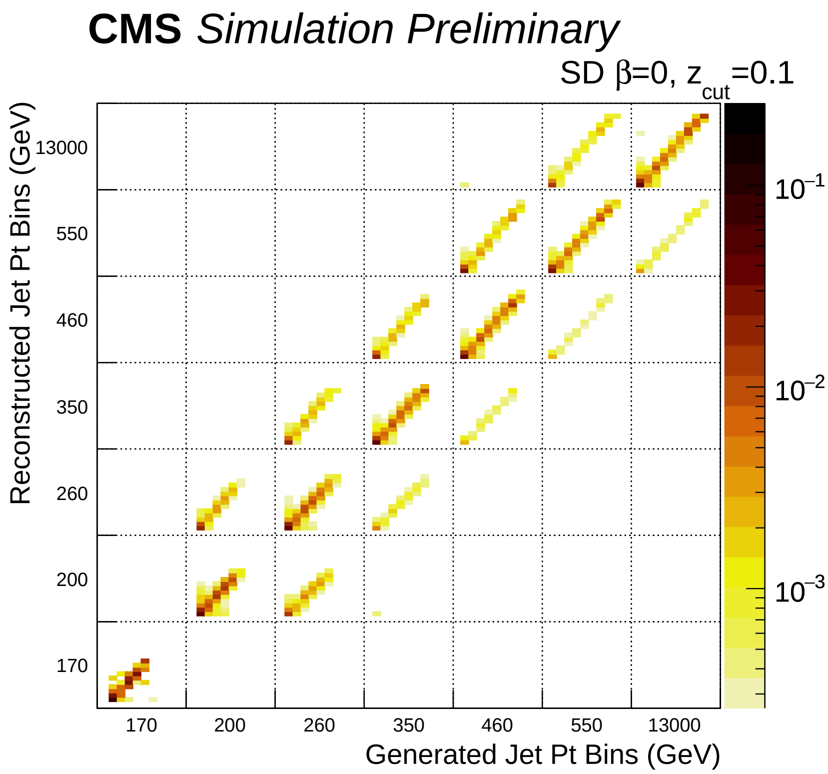

Figure 5.14: Two-dimensional response matrix for groomed jets $\beta = 0$, $z_{cut} = 0.1$. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV).while the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

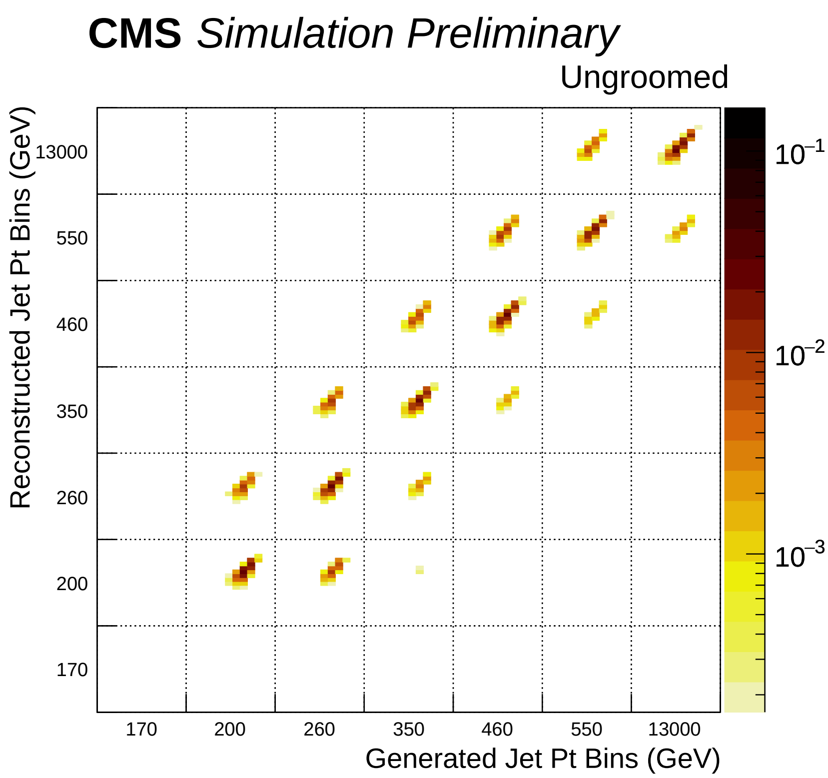

Figure 5.15: Two-dimensional response matrix for ungroomed jets. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV). while the jet dimensionless mass is indexed by the minor blocks on the x (y) axis -5.0, -3.5, -3.0, -2.5, -2.0, -1.5, -1.0, -0.5, 0.0 (-5.0, -4.0, -3.5, -3.25, -3.0, -2.75, -2.5, -2.25, -2.0, -1.75, -1.5, -1.25, -1.0, -0.75, -0.5, -0.25, 0.0).

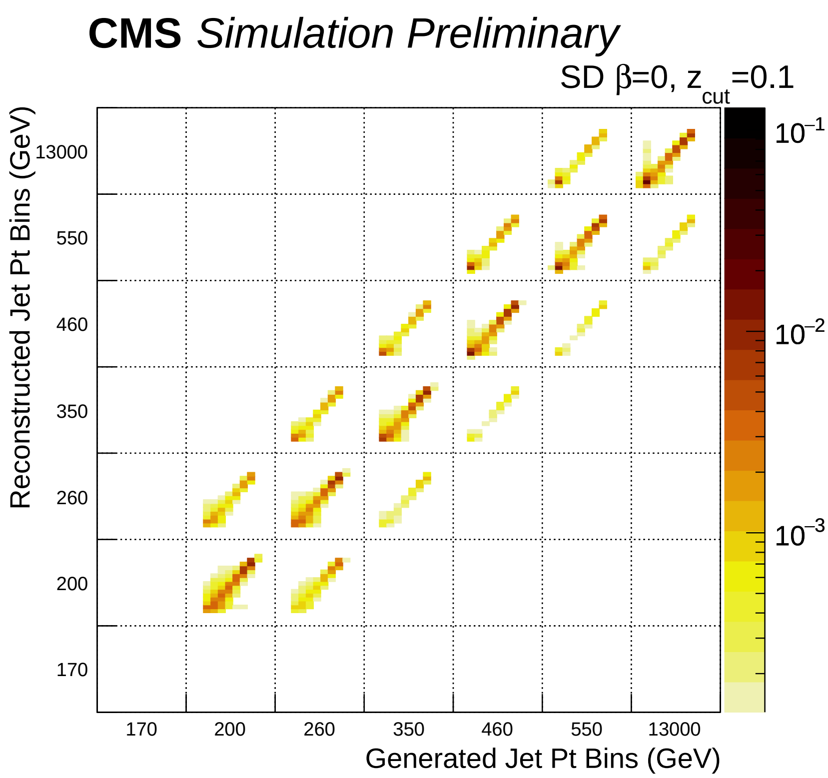

Figure 5.16: Two-dimensional response matrix for groomed jets. The bins are set such that the p_T is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV). while the jet dimensionless mass is indexed by the minor blocks on the x (y) axis -5.0, -3.5, -3.0, -2.5, -2.0, -1.5, -1.0, -0.5, 0.0 (-5.0, -4.0, -3.5, -3.25, -3.0, -2.75, -2.5, -2.25, -2.0, -1.75, -1.5, -1.25, -1.0, -0.75, -0.5, -0.25, 0.0).

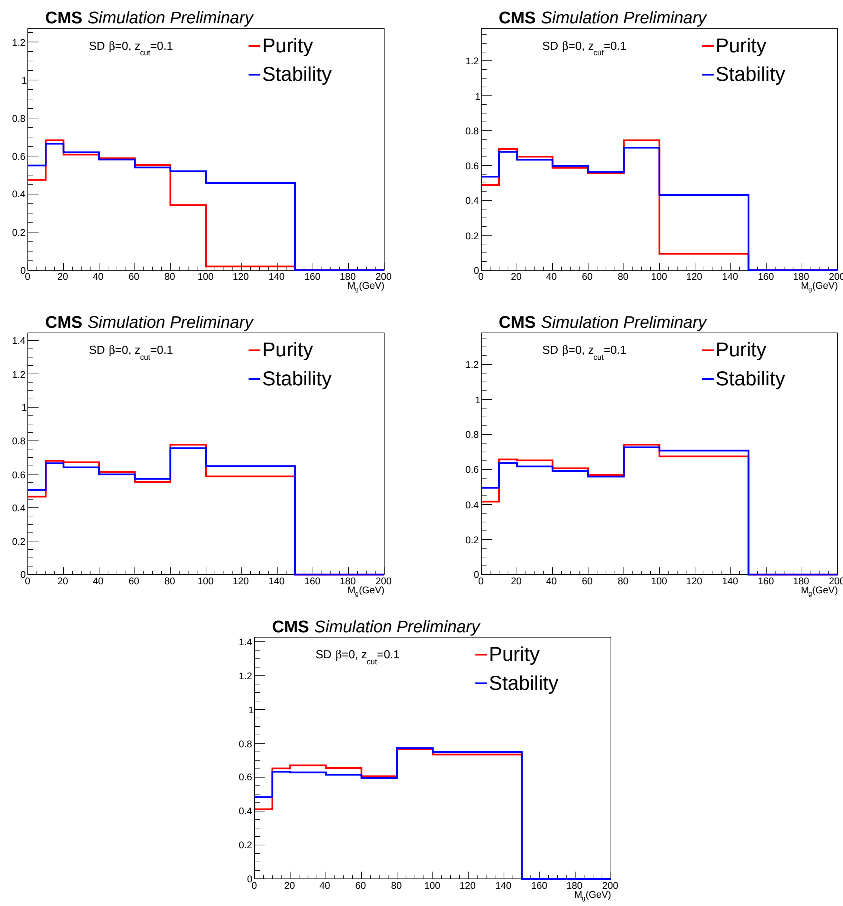

Figure 5.17: Purity and stability for ungroomed jets in various p_{T} bins.

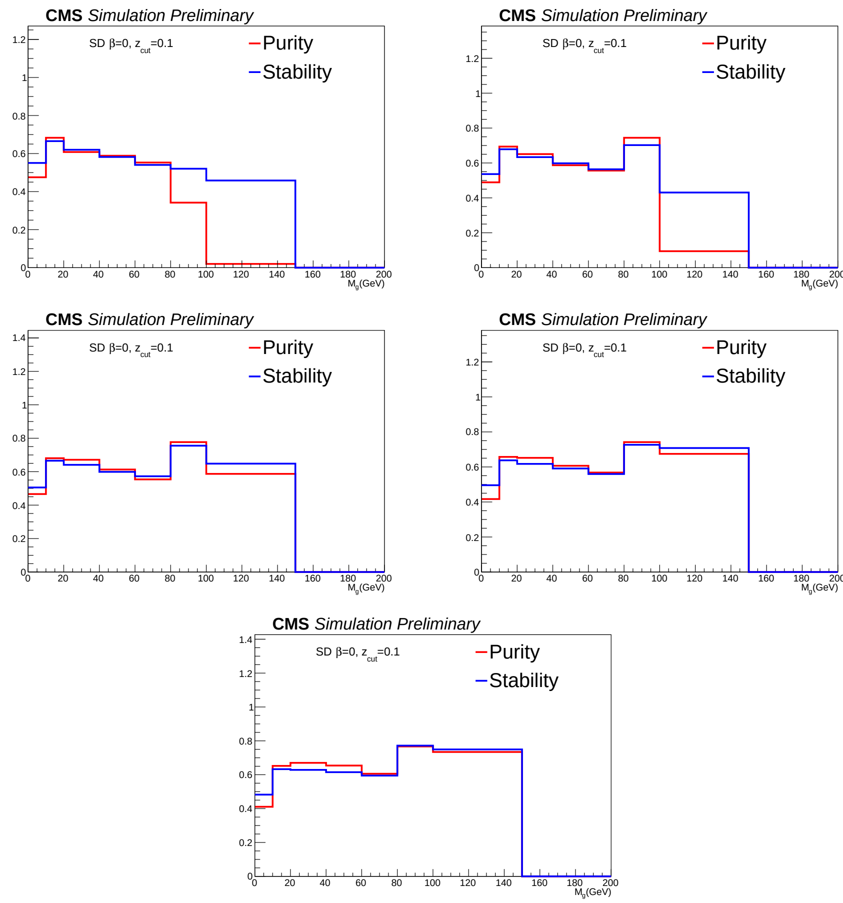

Figure 5.18: Purity and stability for groomed jets, where the soft-drop criterion was applied with $\beta = 0$ and $z_{cut} = 0.1$, in various p_{T} bins.

5.7 Uncertainties

The statistical uncertainties from the MC are non-trivial to estimate due to the unfolding procedure. Instead of standard methods, we employ a "jackknife resampling" technique (a linear bootstrapping method), where we create 10 response matrices that each remove an exclusive 10% of the sample. Each event is removed exactly once. The RMS of this ensemble of distributions, multiplied by 10/9ths, is taken as the statistical uncertainty [47].

There are both theoretical and experimental systematic uncertainties that affect this measurement, along with uncertainties from the unfolding procedure. Unless otherwise specified, the uncertainties are estimated by creating response matrices with the systematic uncertainty applied, and unfolding the reconstructed data or MC. The difference between the systematically varied output and the nominal output is taken to be the systematic uncertainty of the measurement. The response matrix uncertainties are symmetrized.

For the normalized cross section, the differences are taken after normalizing for each individual p_{T} bin. This removes much of the p_{T} dependence of the measurement since it cancels in the ratio. For the full cross section, the differences are taken without normalizing each p_{T} bin, so include these effects. As such, the normalized cross section has very small jet energy uncertainties, whereas the absolute cross section has large jet energy uncertainties. The luminosity uncertainty also cancels in the ratio.

The predominant experimental uncertainties that affect the normalized cross section measurement are the jet mass scale, the jet mass resolution, the parton shower uncertainty, the pileup reweighting scheme, the jet energy scale, and the jet energy resolution. The predominant theoretical uncertainty is the physics model.

In addition to the previous effects, the absolute cross section also has sizeable uncertainties from the jet energy scale, jet energy resolution, and overall luminosity measurement.

- **Jet energy scale**: The JES uncertainty is evaluated by varying the `L1L2L3` corrections up and down by the appropriate uncertainties as described in Ref [**?**].

- **Jet energy resolution**: The JER uncertainty is evaluated by smearing the jet energy by an additional factor as described in Ref [**?**]. The central value is determined by smearing to the JER nominal value, and the uncertainties are evaluated by smearing to the JER systematic variations.

- **Jet mass scale**: The JMS in Run 1 was equal to unity within experimental uncertainties, derived from investigating the mass peak of the W boson within $t\bar{t}$ events, and comparing the results in data and MC. In Run 2, this is still effectively unity within statistical uncertainties [48]. The uncertainty on this is less than 1%, and is thus neglected.

- **Jet mass resolution**: The JMR in Run 1 was equal to 1.11 ± 0.09. In Run 2 we now observe 1.07 ± 0.12. The JMR is evaluated by stretching the jet mass by 7% relative to the generator-level jets [48]. The uncertainty is set to be the uncertainty in this measurement, 12%.

- **Pileup**: The pileup uncertainty is estimated by reweighting the MC, assuming the minimum bias cross section of 69 mb is scaled up and down by 1.4%.

- **Luminosity**: The CMS measurement of the luminosity for the 2015 data-taking period is 2.7% [49].

- **Parton distribution functions**: The PDF uncertainty is evaluated by using the `PDF4LHC15` meta-PDF set.

- **Physics model uncertainty**: The PS uncertainty, which includes the parton shower (PS), generator tune, and physics showering model, is evaluated by comparing a response matrix derived from PYTHIA8 to a response matrix derived from HERWIG++. The difference between the response matrices for the two cases is taken as a shifted systematic uncertainty.

Figure-5.19 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_{T} bin for the normalized cross section.

Figure-5.20 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_{T} bin for the normalized cross section.

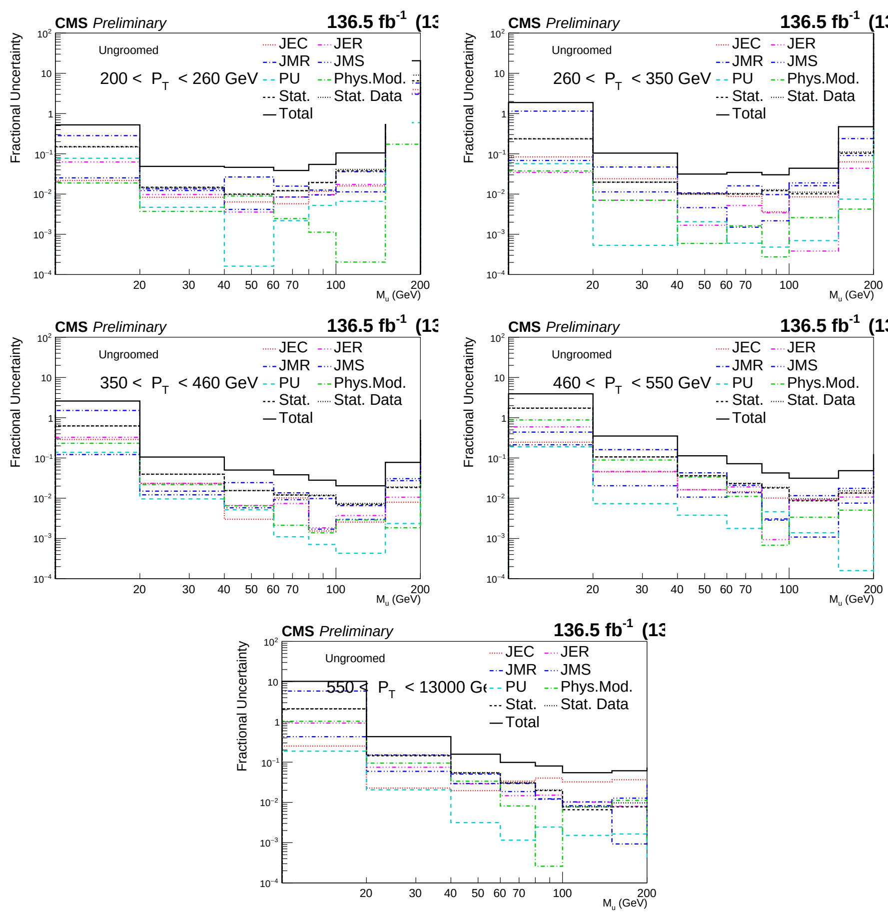

Figure 5.19: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the raw jet mass.

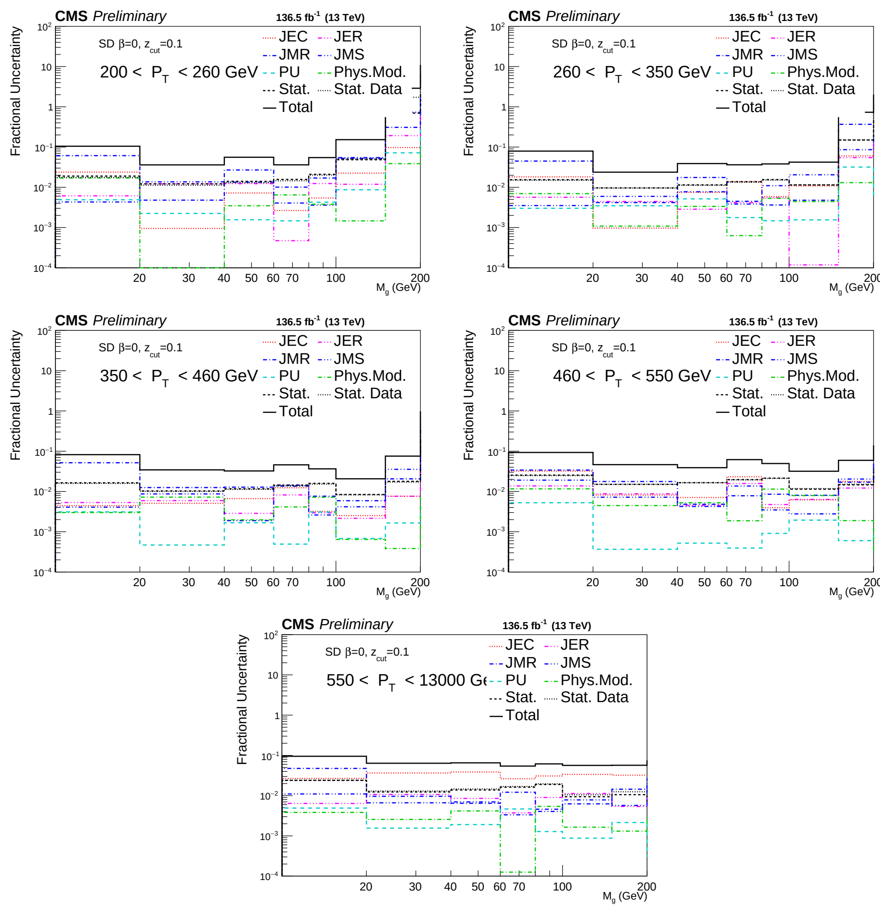

Figure 5.20: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_{T} bins for the jet mass after grooming with default Soft-Drop.

5.8 Unfolding

The unfolding procedure in this analysis is a least squares fitting with optional Tikinov regularization, implemented using the TUnfold algorithm [50]. This procedure was used rather than the D'Agostini [51] method employed in previous measurements as it has the capacity to accept variant truth MC vectors as input, not simply the same one used to fill the response matrix as required in the current RooUnfold version.This issue was brought to our attention due to the recommendation of the CMS statistics committee and their claim that TUnfold is better documented. We have found this claim to be unfounded in the two dimensional case however we have developed some informational materials outlining TUnfold in 2D using the algorithm's built-in systematic uncertainty handling. Regardless, moving to TUnfold from RooUnfold prevents a bias towards the input distribution since TUnfold allows the option of using different input MC truth vectors, not necessarily filled with the same events as the response matrix, which is lacking in RooUnfold and causing regularization to bias the results towards the input distributions. This is currently the first analysis that we are aware of on CMS, which makes use of TUnfold in 2D using the built-in systematic uncertainty handling.

The response matrix was used to unfold the MADGRAPH + PYTHIA8 MC with and without Tikhinov regularization .It was determined that regularization is not necessary for the current binning scheme due to the high purity and stability of the binning scheme. Furthermore, the condition number was determined to be 1.43 which was considered small enough, less than 10 is suggested by the statistics committee currently, such that regularization was not required. A simple matrix multiplication was then used to re-fold the unfolded MC without regularization. The resulting re-folded distribution was compared to the original reconstructed distribution.

Several validations were performed to ensure that the unfolding procedure was accurately reproducing input distributions.

1. **Closure test in MC**: The PYTHIA8 MC was used to construct both the response matrix as well as the reconstructed distribution. By construction, this should reproduce exactly the input given, since that is where it was derived from. Figure **??** showing the closure tests for ungroomed jets for all p_{T} bins, and Figures 5.31-5.34 show the same for groomed jets.

2. **Physical bias test in MC**: To investigate the capability of the response matrix to correctly unfold a spectrum different from the inputs, we have performed another closure test unfolding the 2016 inclusive HERWIG++ sample as the input distribution with the 2016 inclusive PYTHIA8 response matrix . This is not the same as the systematic variation due to the parton shower uncertainty, where the difference between the `PYTHIA8` response matrix and the HERWIG++ response matrix is treated as a shifted uncertainty, using TUnfold's internal systematic uncertainty handling, applied to the unfolding of the nominal PYTHIA8 reconstructed response matrix.

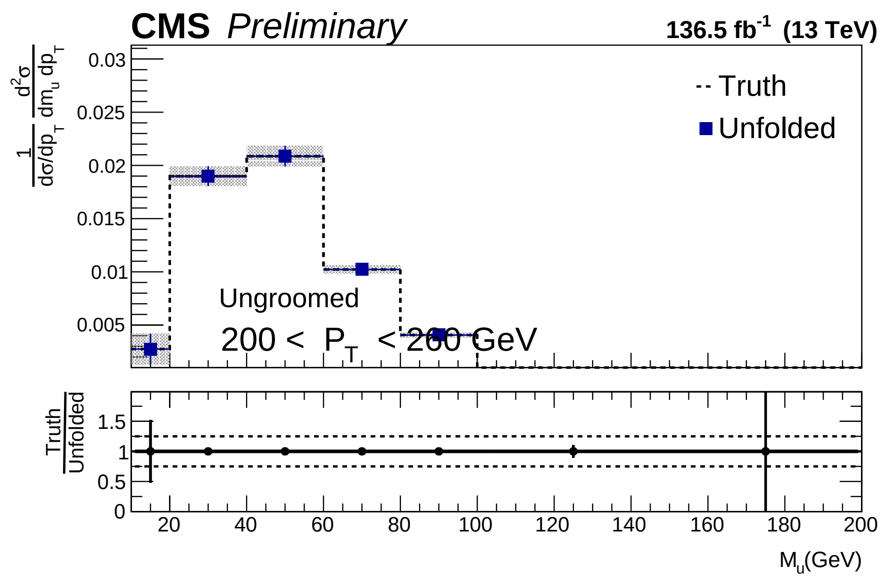

Figure 5.21: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

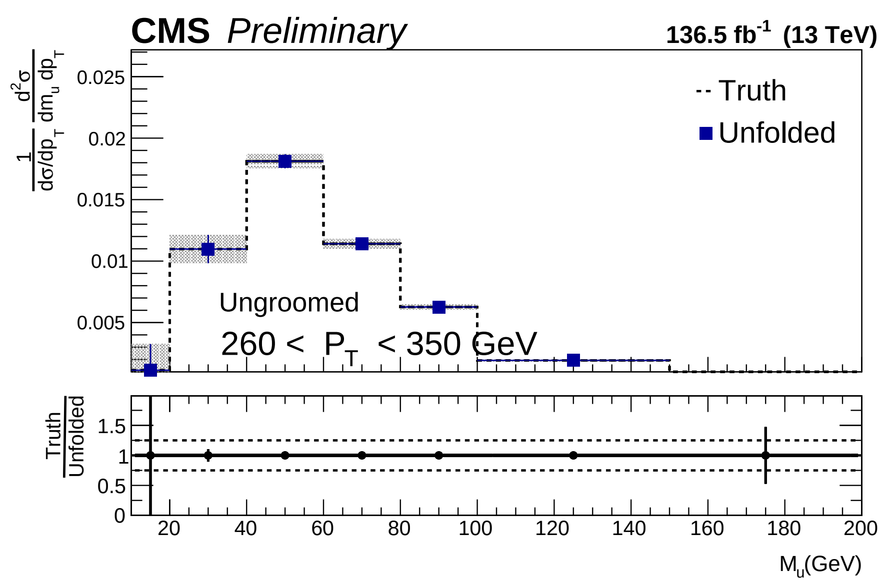

Figure 5.22: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

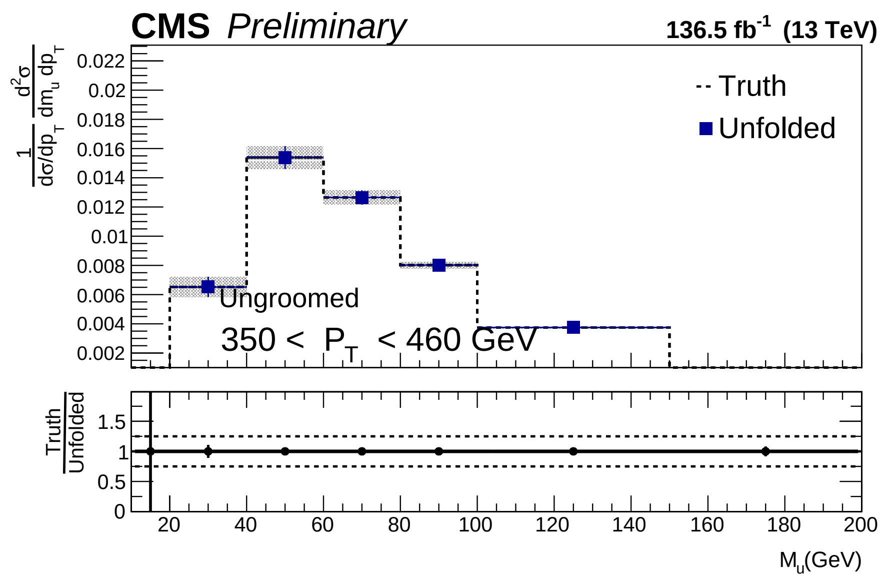

Figure 5.23: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

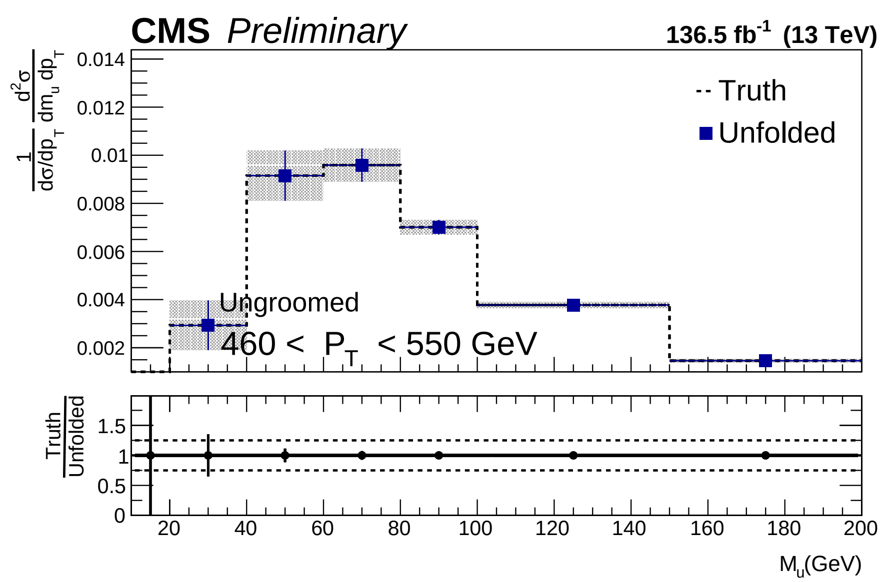

Figure 5.24: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

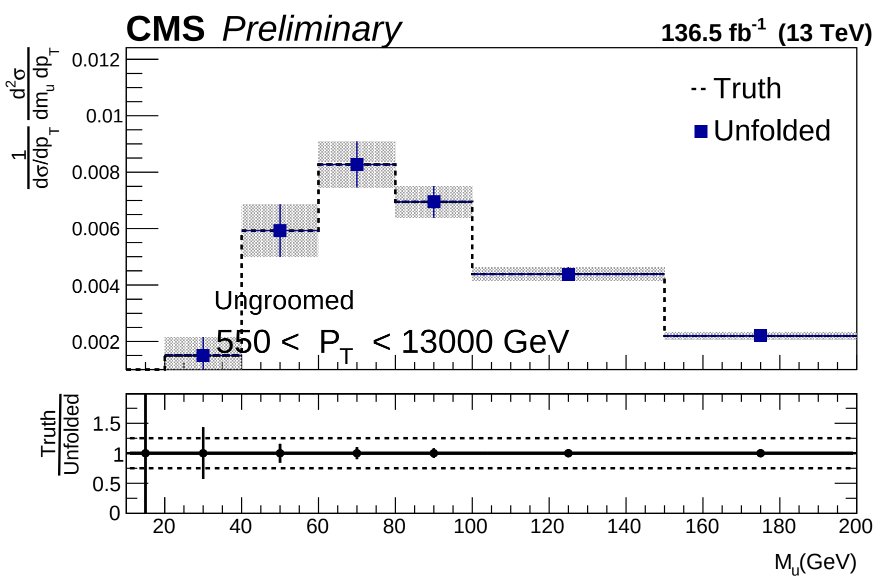

Figure 5.25: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

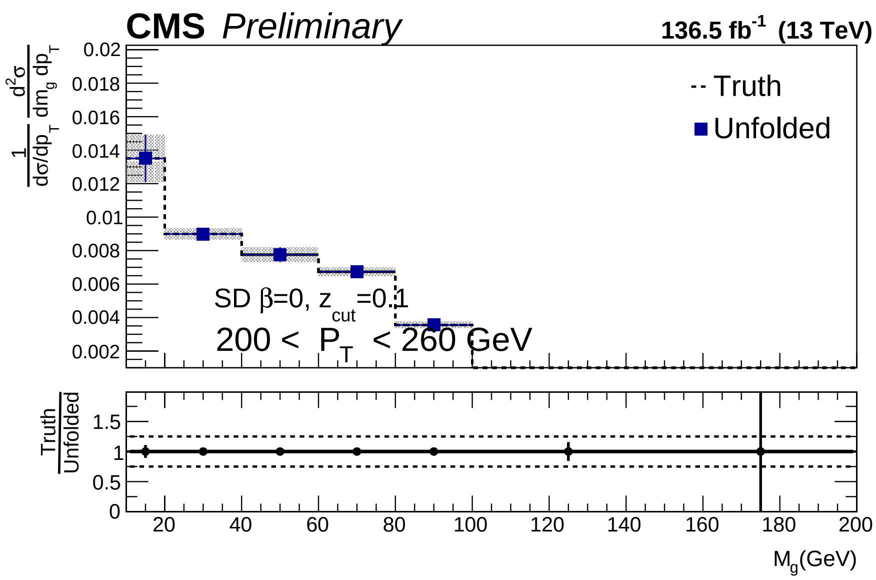

Figure 5.26: Closure test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

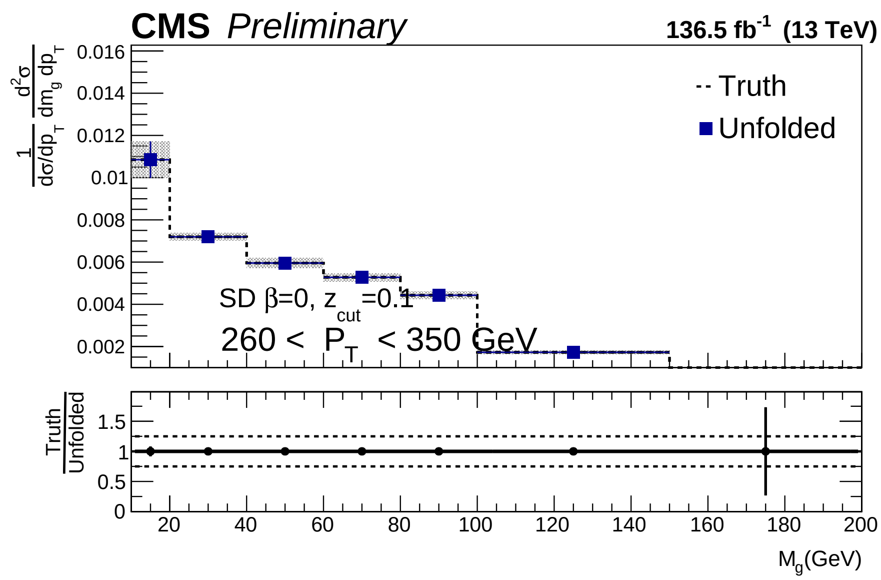

Figure 5.27: Closure test of groomed reconstructed Monte Carlo, p_T 260-350 GeV.

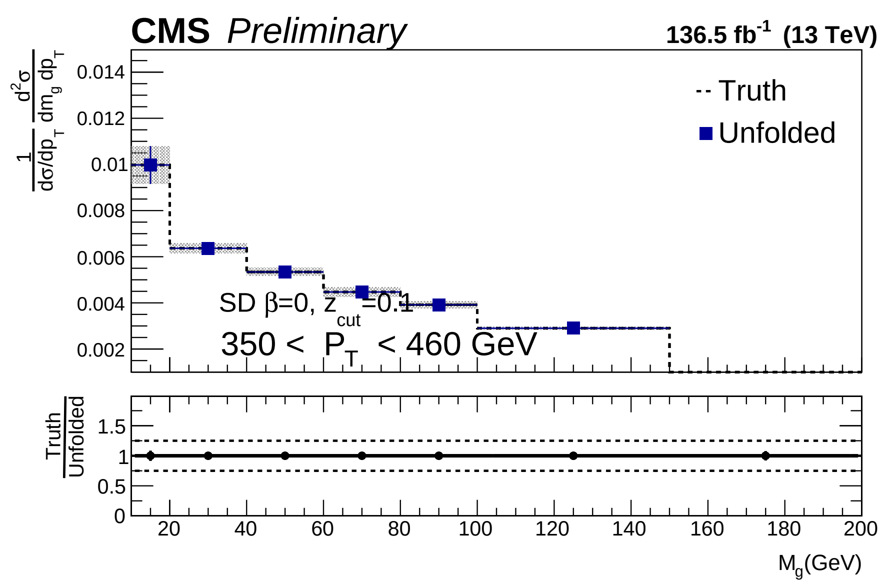

Figure 5.28: Closure test of groomed reconstructed Monte Carlo, p_T 350-460 GeV.

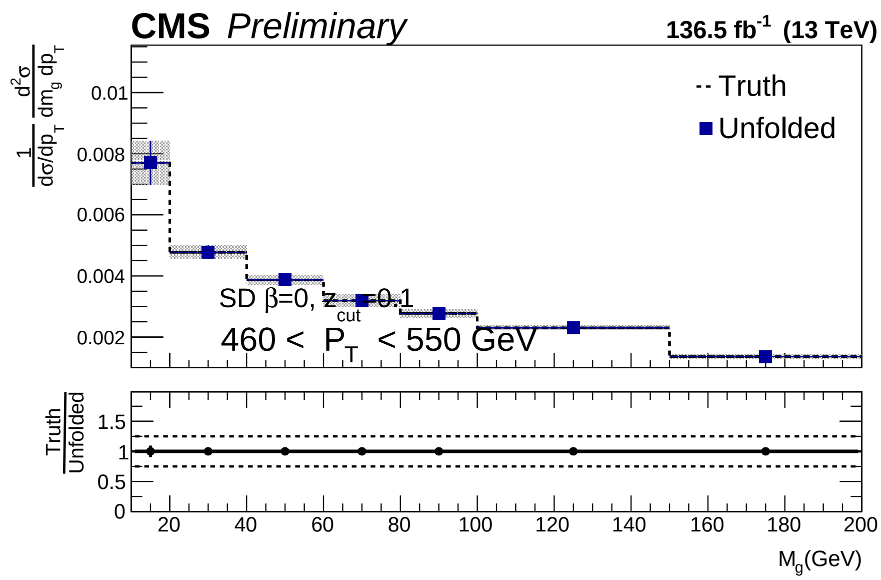

Figure 5.29: Closure test of groomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

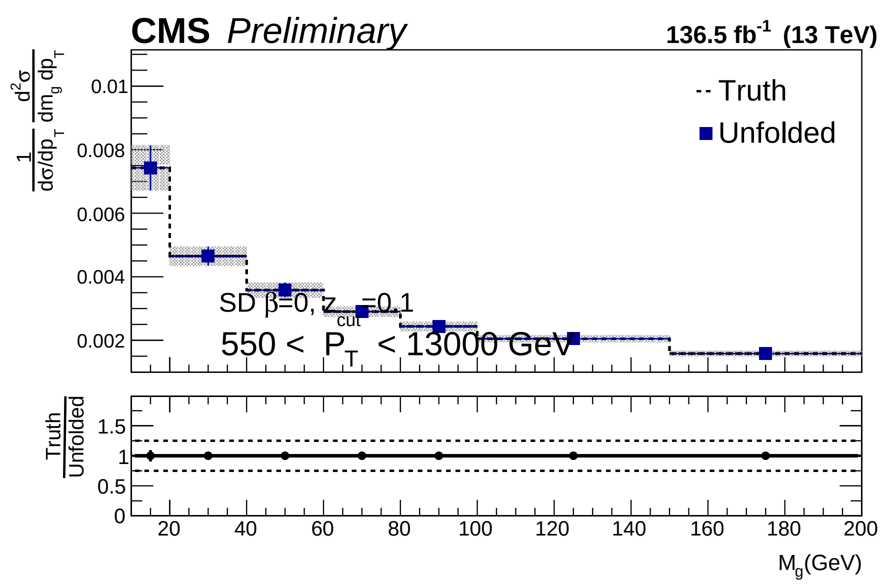

Figure 5.30: Closure test of groomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

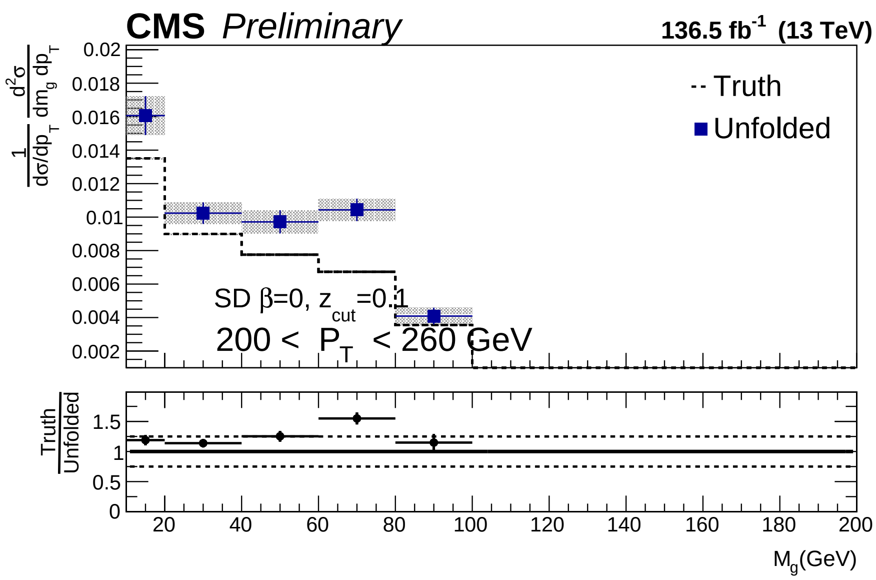

Figure 5.31: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

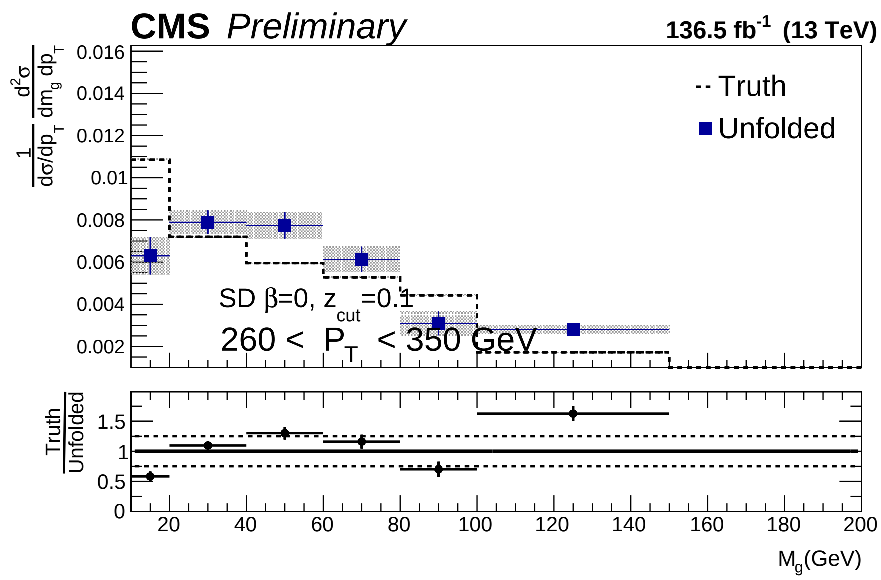

Figure 5.32: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

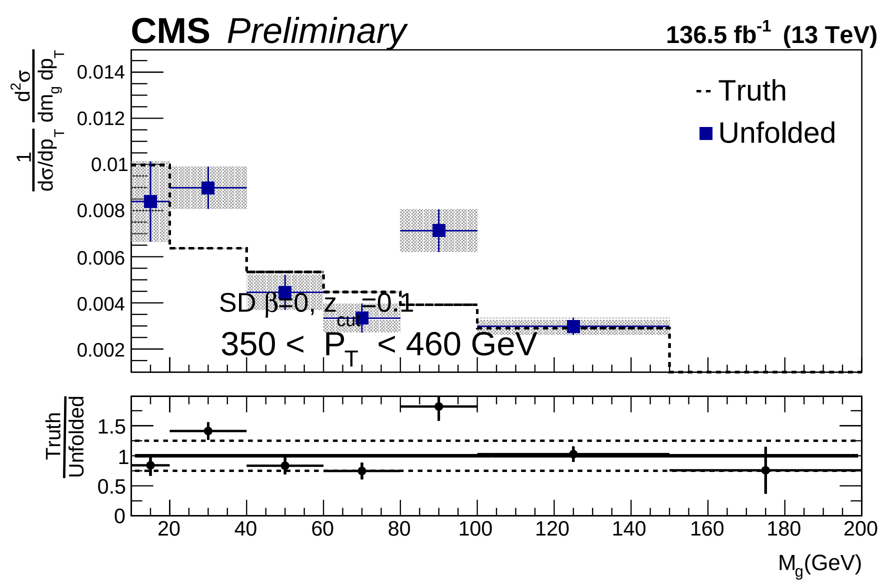

Figure 5.33: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

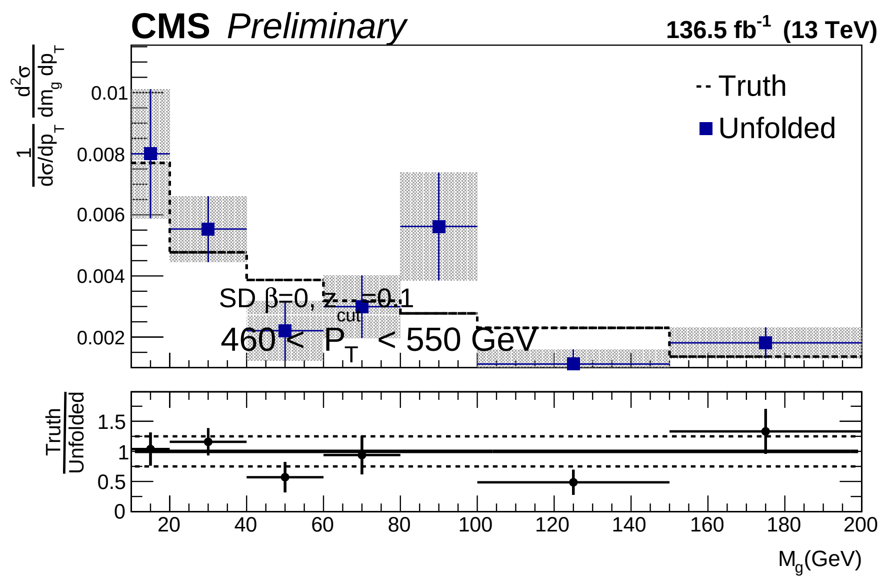

Figure 5.34: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

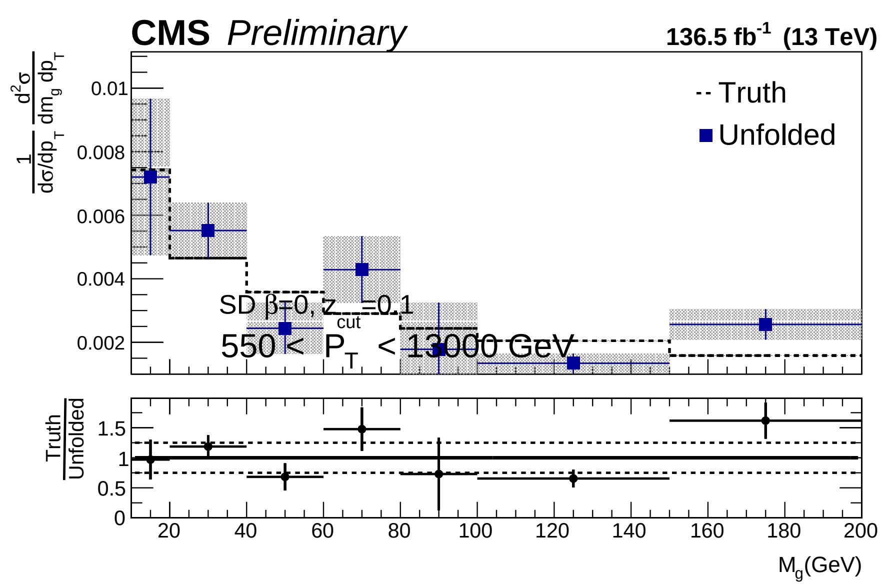

Figure 5.35: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

5.9 Results

The unfolded jet p_T spectrum is shown in Fig. **??**. The dominant uncertainty is the jet energy scale.

The normalized cross section results for the ungroomed jets are shown in Figures **??**-**??**, and for the groomed jets are shown in Figures **??**-**??**. The distributions are shown normalized per p_T bin to compare only the shapes. All p_T bins are shown in double-logarithmic plots in Figures 5.46 and 5.47.

In order to visualize the results in a slightly different way, we also present the normalized cross section with respect to $log(m)$. This is equivalent to multiplying the unfolded cross section by the mass. We are approximating this by multiplying each bin in the final result by the average mass in that bin. Thus, we present

$$\frac{m}{d\sigma/dp_T}\frac{d^2\sigma}{dm\,dp_T}. \tag{5.3}$$

These results are shown for ungroomed jets in Figures **??**-**??**, and for the groomed jets are shown in Figures 5.41-5.43. The distributions are normalized per p_T bin and multiplied by the center of the mass bin.

The results are currently presented compared to MADGRAPH + PYTHIA8 and will soon be compared to HERWIG++ when the 2017 samples are ready, and, for the normalized groomed case, we will eventually compare the calculation of the jet mass from Marzani et al. in progress now, expected in early 2020. The normalized results are scaled to unity for each individual jet p_T bin for the unfolded data, the MADGRAPH + PYTHIA8, and results from the calculation from the authors of Ref. [20]. The calculation from Ref. [32], on the other hand, is normalized to match the data at 50 GeV, since the calculation does not consider nonperturbative corrections that are important below that value. Furthermore, at very high masses (relative to the p_{T}), the LO matrix element matching is also insufficient to capture the true dynamics, since at LO there is a kinematic turn-off at $p_{\mathrm{T}}R/2$ (where R is the distance parameter for the jet clustering), from the relativistic kinematics of a $1 \rightarrow 2$ decay. However, for real jets the turn-off is closer to $p_{\mathrm{T}}R/\sqrt{2}$ due to stochastic effects. To see this, consider a particle of energy E and mass m decaying to two massless particles, each with an energy $E/2$ and separated by an angle θ. The mass must satisfy $m^2 < \frac{E^2}{2}(1 - \cos\theta)$. In the small angle limit, this would be $m^2 < E^2\theta^2/4$, or $m < E\theta/2$. With more particles, the stochastic nature of the shower increases this to $m < E\theta/\sqrt{2}$. Thus, a leading-order ($1 \rightarrow 2$) decay will have a faster kinematic turn-off than an all-orders ($1 \rightarrow$ many) decay.

The data are reproduced reasonably well by the simulations, with poorer agreement below the Sudakov peak. The matrix element does not have a large impact on the results due to the excellent agreement between `Madgraph+PYTHIA8` and `HERWIG++` alone, however the parton shower program chosen can have larger effects. Values of the χ^2 probabilities for the data-to-MC and data-to-theory comparisons are shown in Tables **??** and **??**. Table **??** shows the probabilities if we restrict the range of the comparison to accommodate known deficiencies in the predictions. At very low masses, nonperturbative effects are large, and at very large masses, the fixed order matching is insufficient to capture the kinematics. As such, the comparisons are restricted to a range of intermediate jet masses. In Table **??** the ranges are unrestricted, for comparison. In addition to these effects, at very low groomed masses, the resolution of the detector itself spoils the agreement between simulation and theory. For these reasons, the probability after the range restriction is more appropriate for comparisons.

The soft drop algorithm considerably lowers the jet mass distribution, as observed in Figs. **??**-**??** and Fig. 5.47. The soft and collinear parts of the jet are removed, leaving the hard jet function, which tends to have low mass. The precision improves after the grooming algorithm is applied, and the physics modeling (primarily parton shower) uncertainty is reduced.

It is also worth noting that the groomed jet distributions exhibit falling spectra for nearly all jet p_T bins, whereas the ungroomed jets have a peaked structure, making unfolding more difficult. The groomed jets are able to be nicely unfolded with much lower uncertainties. In addition to this feature, the parton shower uncertainty, which is one of the largest uncertainties for the ungroomed jets below the Sudokov peak, is dramatically reduced after grooming.

Checks on the unfolding are shown in Figures 5.52-**??**. Figures 5.52-**??** and Figures 5.57-**??** show the ratio of the unfolded to raw data and MC for ungroomed and groomed jets, respectively. Figures 5.62-**??** and Figures 5.67-**??** show the generator truth over the unfolded data, divided by the raw MC over the raw data. The behaviors observed are well predicted by the MC. There are some features at very low masses (below the Sudokov peak) that the unfolding introduces, but this is reproduced by the MC.

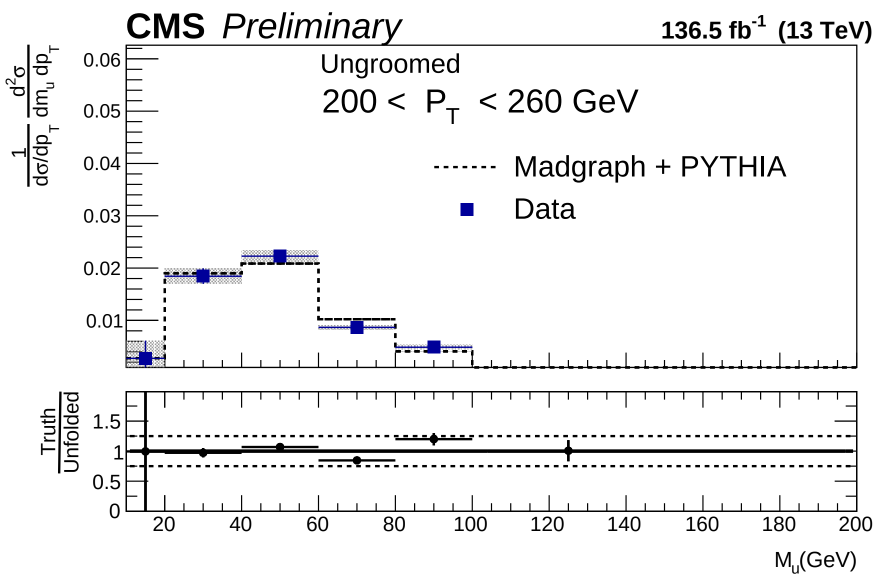

Figure 5.36: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 200-260 GeV.

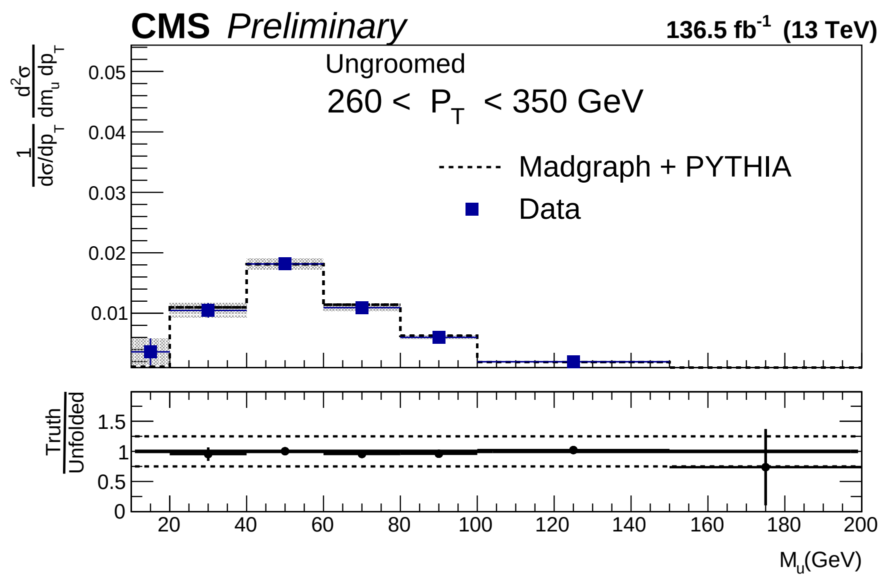

Figure 5.37: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 260-350 GeV.

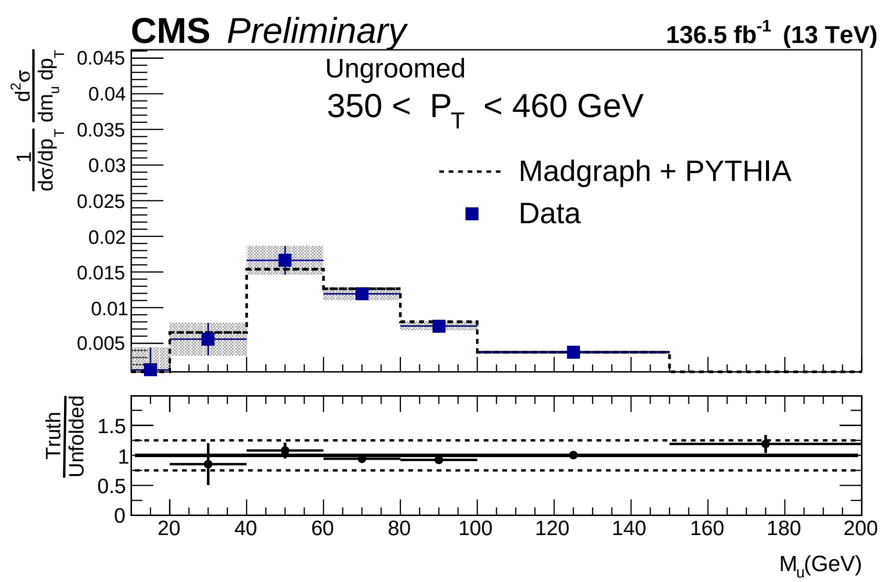

Figure 5.38: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 350-460 GeV.

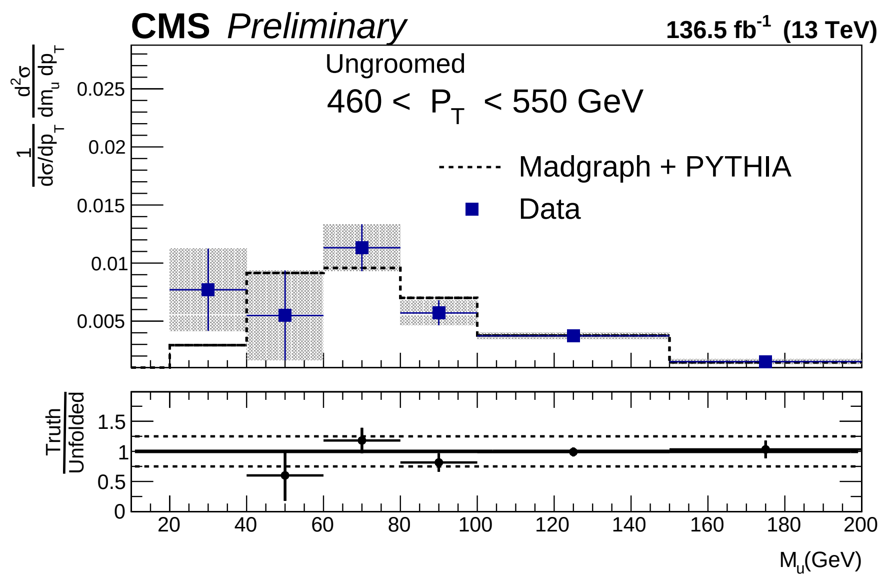

Figure 5.39: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 460-550 GeV.

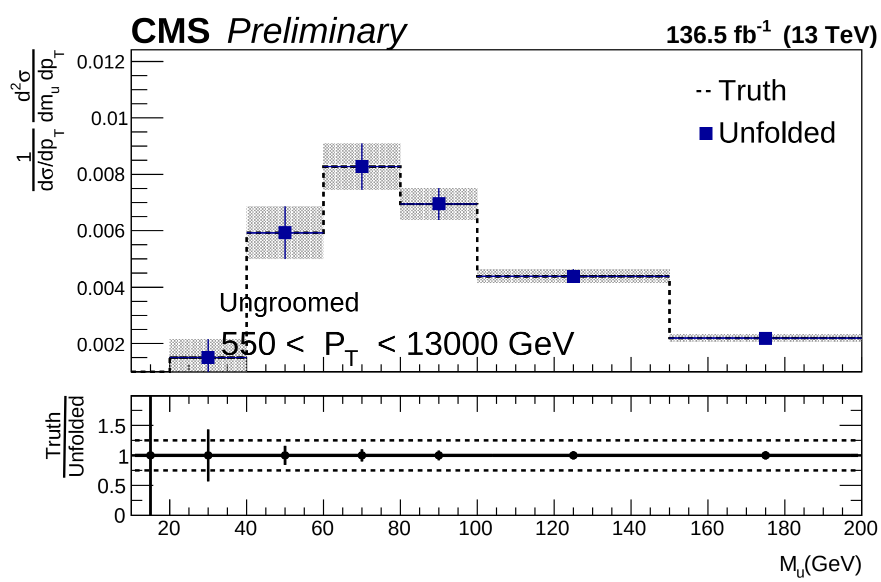

Figure 5.40: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 550-13000 GeV.

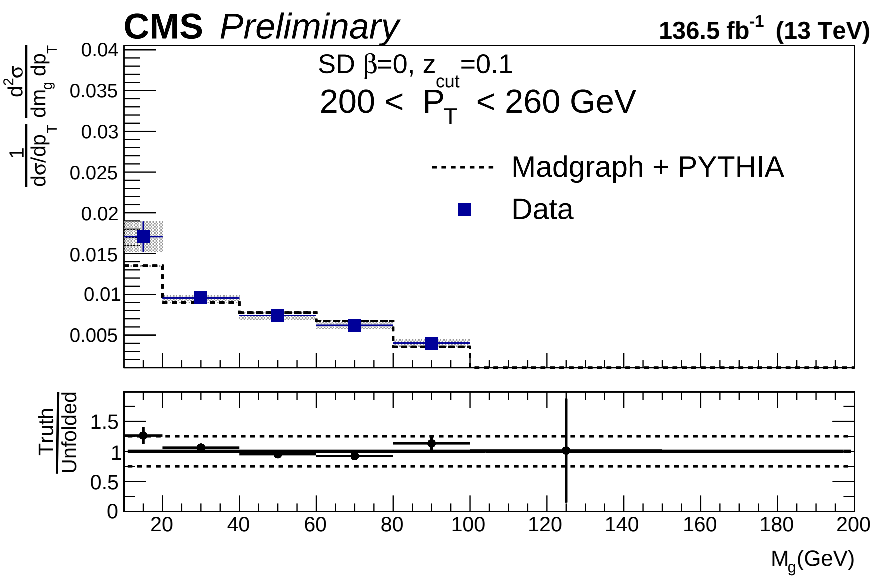

Figure 5.41: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 200-260 GeV.

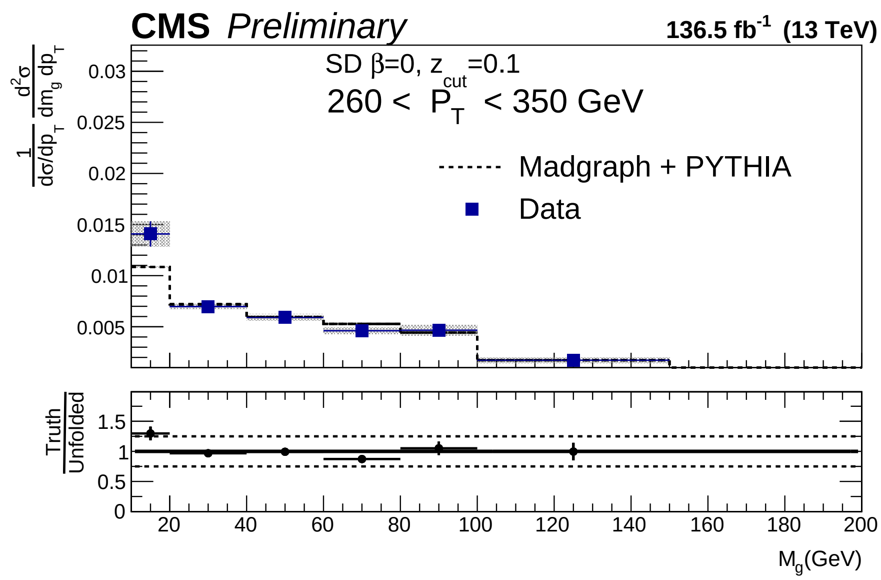

Figure 5.42: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 260-350 GeV.

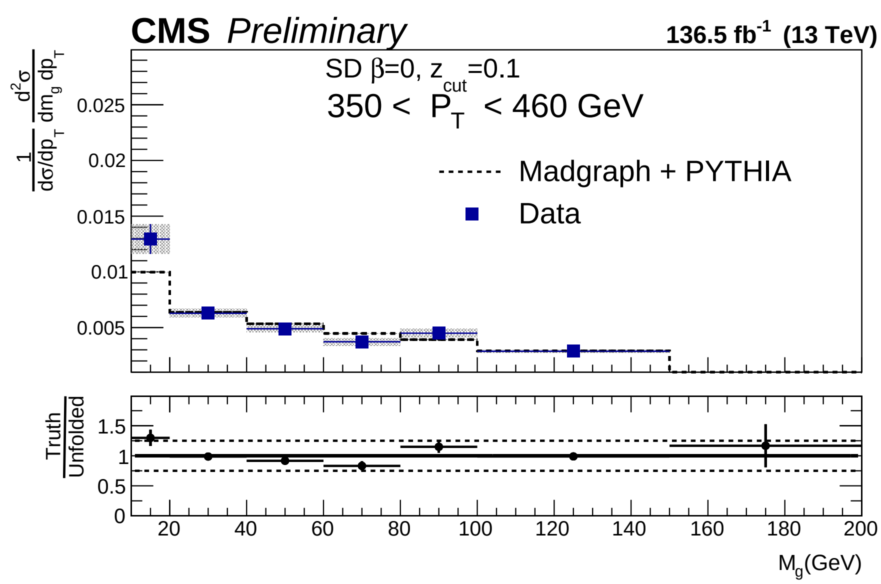

Figure 5.43: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 350-460 GeV.

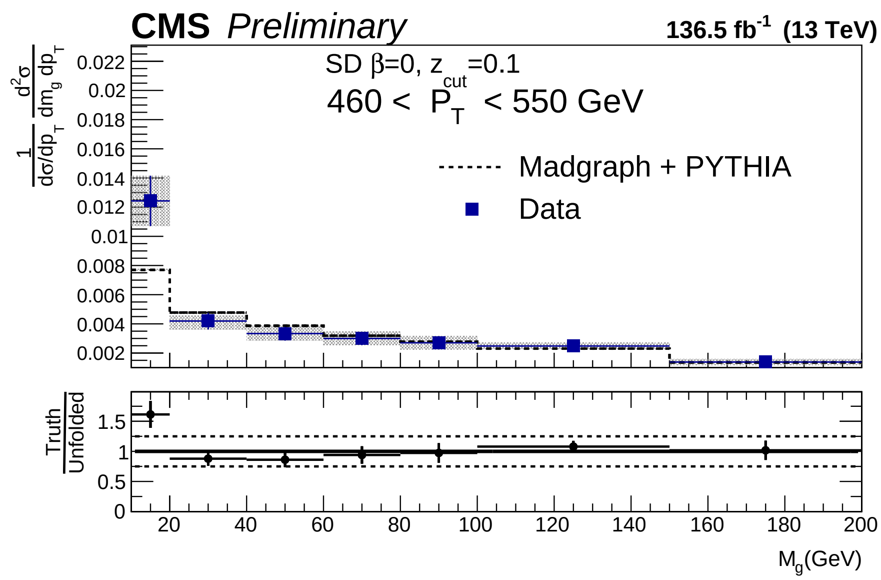

Figure 5.44: Normalized cross section results with respect to jet mass for groomed jets, p_T 460-550 GeV.

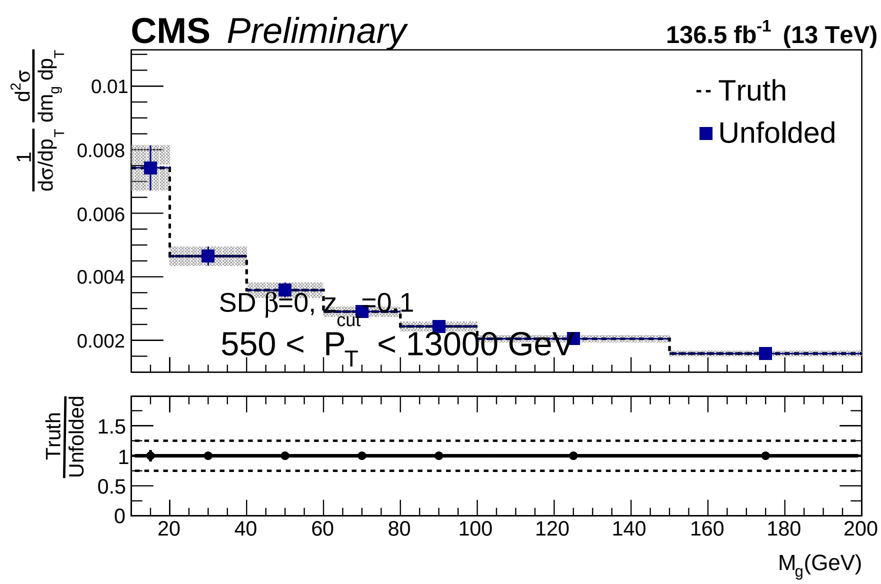

Figure 5.45: Normalized cross section results with respect to jet mass for groomed jets, p_T 550-13000 GeV.

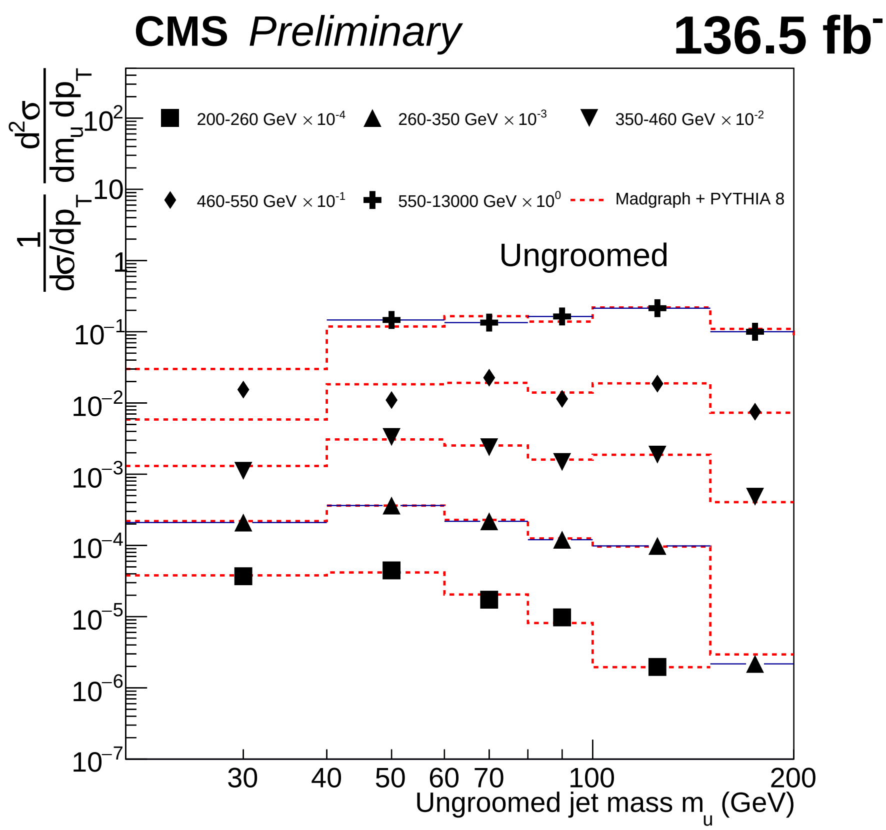

Figure 5.46: Results for ungroomed reconstructed unfolding with jet mass.

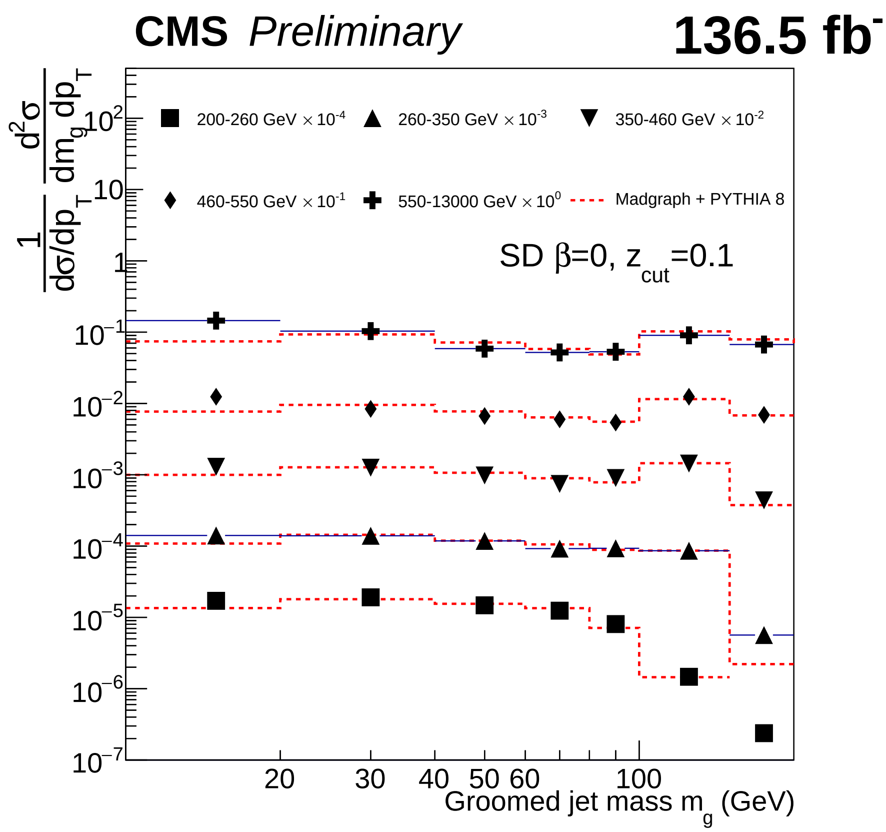

Figure 5.47: Results for groomed reconstructed unfolding with jet mass.

The resulting correlation matrix without systematic uncertainties for the ungroomed jets is shown in Fig. 5.48, and for the groomed jets is shown in Fig. 5.49. The same figures with systematic uncertainties are shown for the ungroomed jets in Fig. 5.50, and for the groomed jets is shown in Fig. 5.51.

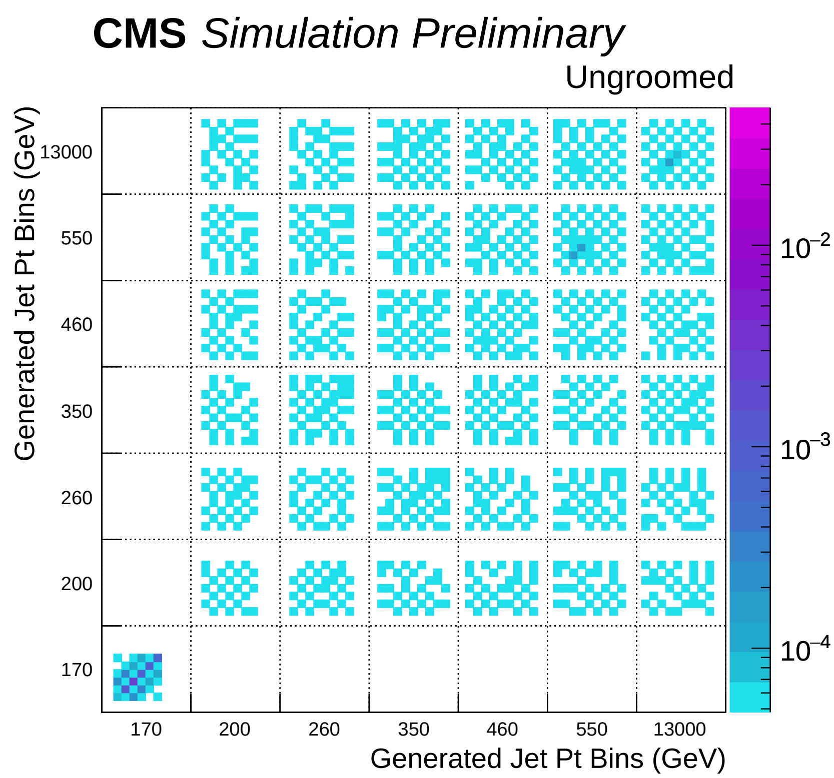

Figure 5.48: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,10,20,. . . 200 GeV), while the p_T is indexed by the major blocks (200,350,. . . 13000 GeV).

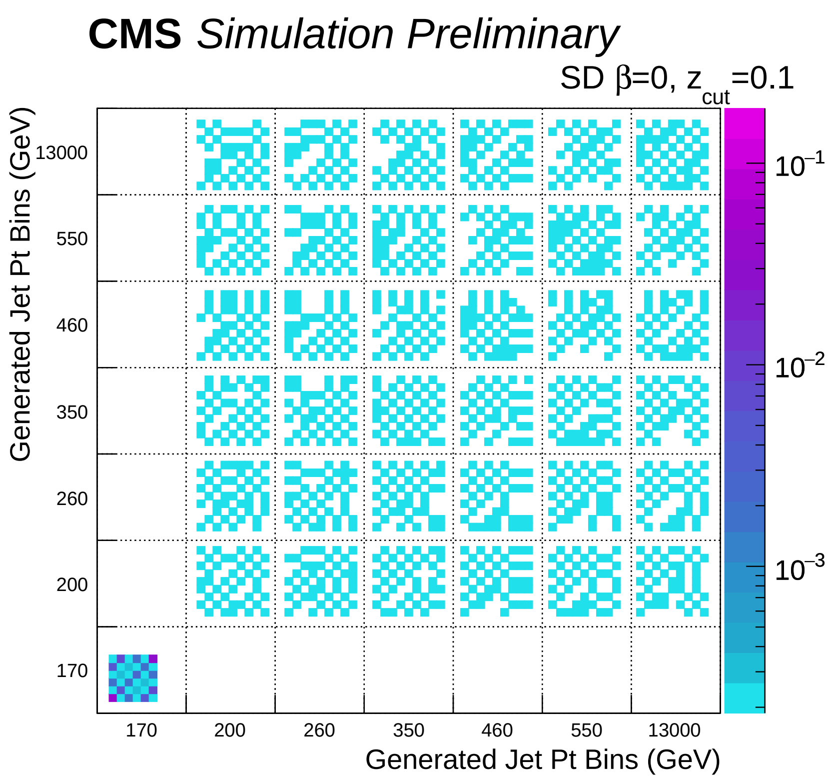

Figure 5.49: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 760 GeV).

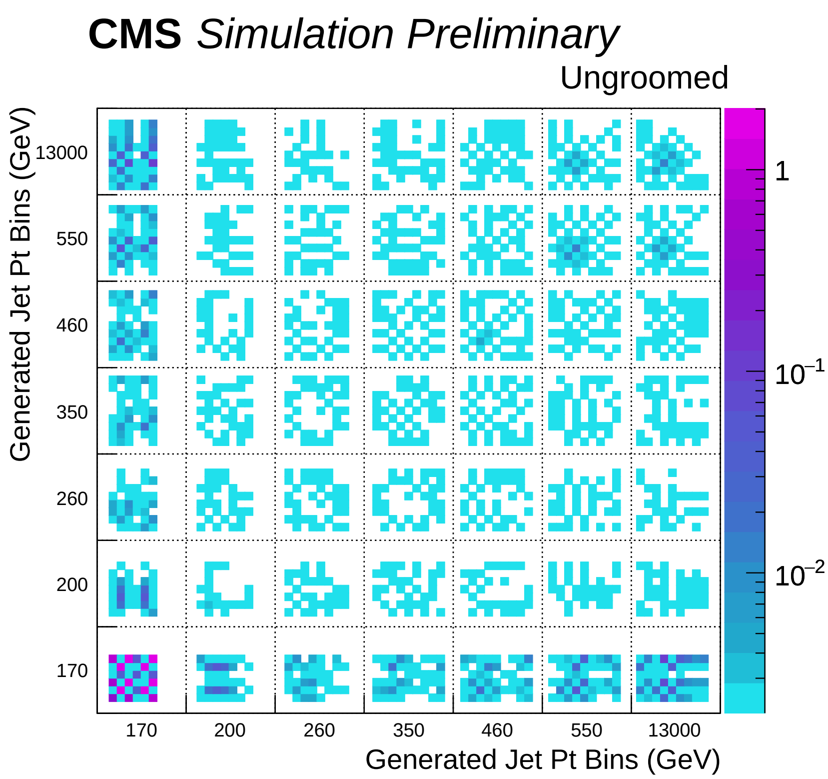

Figure 5.50: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,260,. . . 13000 GeV).

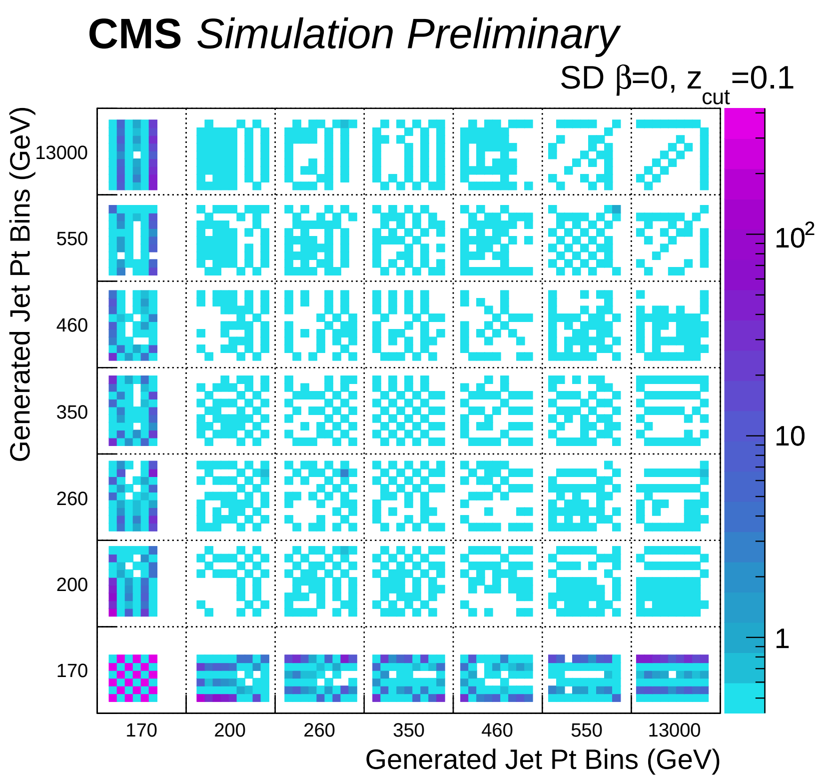

Figure 5.51: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 0 GeV).

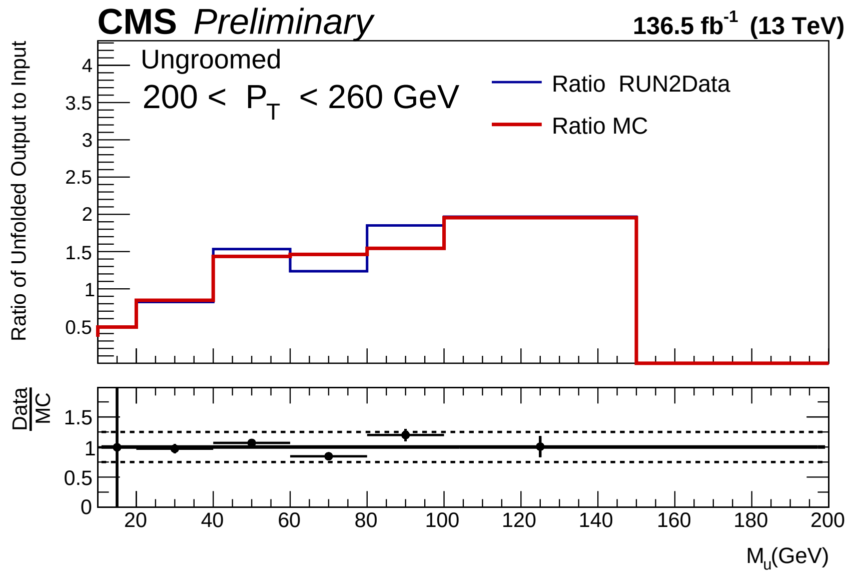

Figure 5.52: Ratio of unfolded over raw data and MC for ungroomed jets, p_{T} 200-260 GeV.

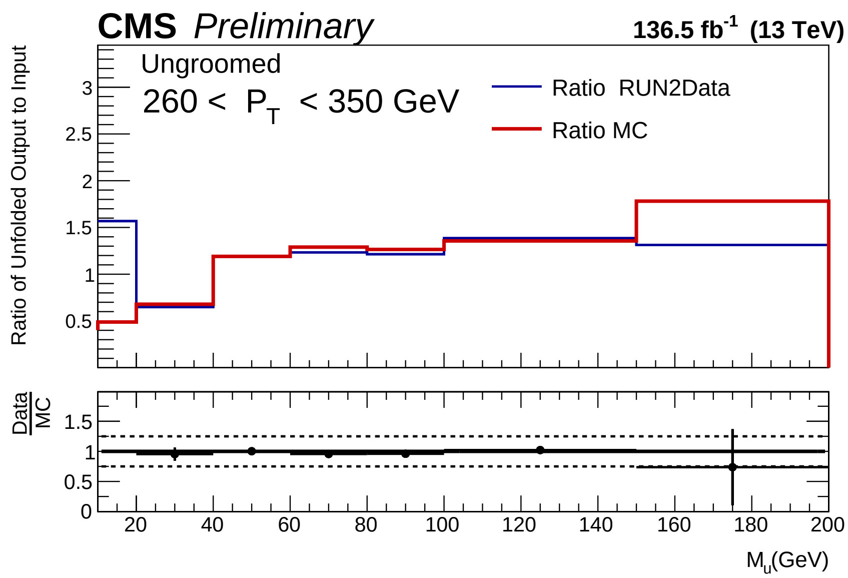

Figure 5.53: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 260-350 GeV.

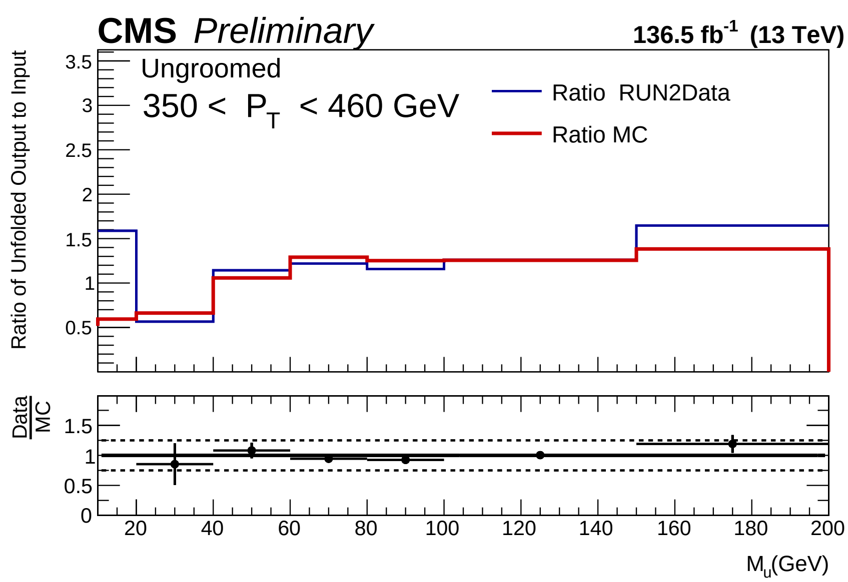

Figure 5.54: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 350-460 GeV.

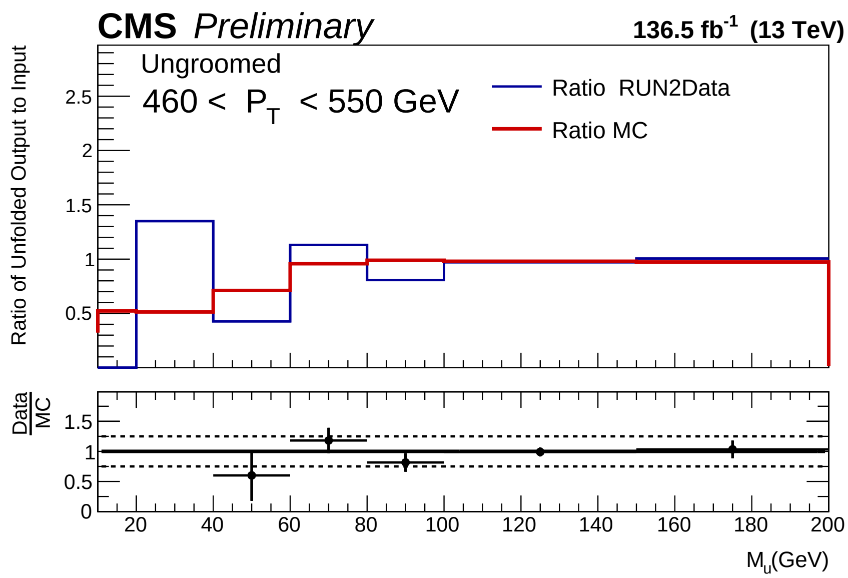

Figure 5.55: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 460-550 GeV.

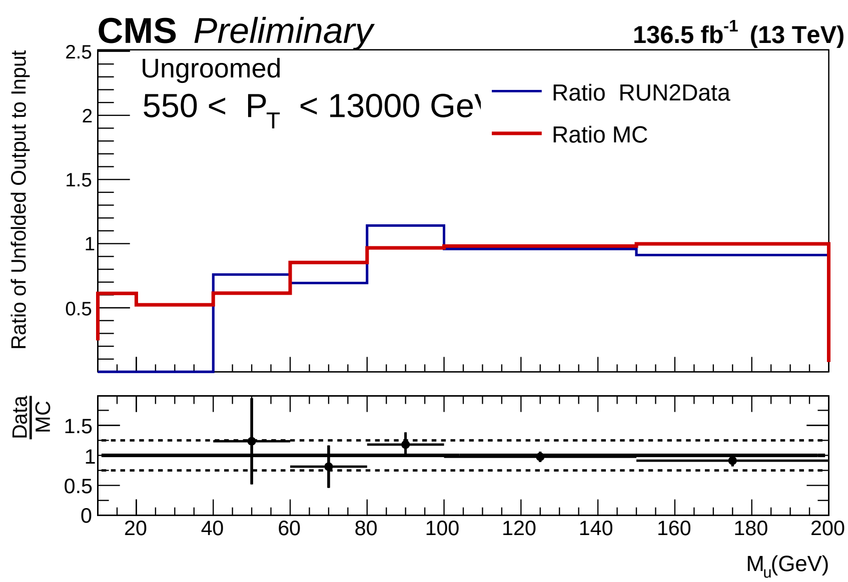

Figure 5.56: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 550-13000 GeV.

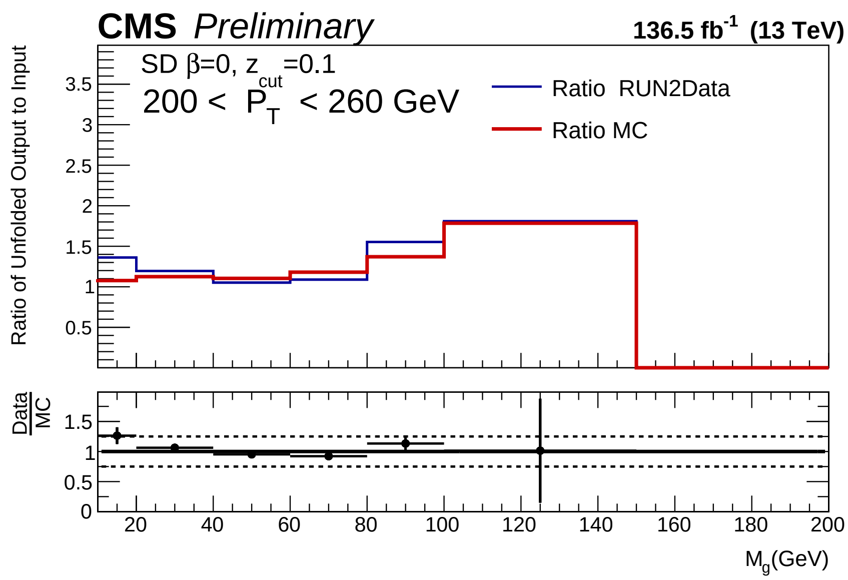

Figure 5.57: Ratio of unfolded over raw data and MC for groomed jets, p_T 200-260 GeV.

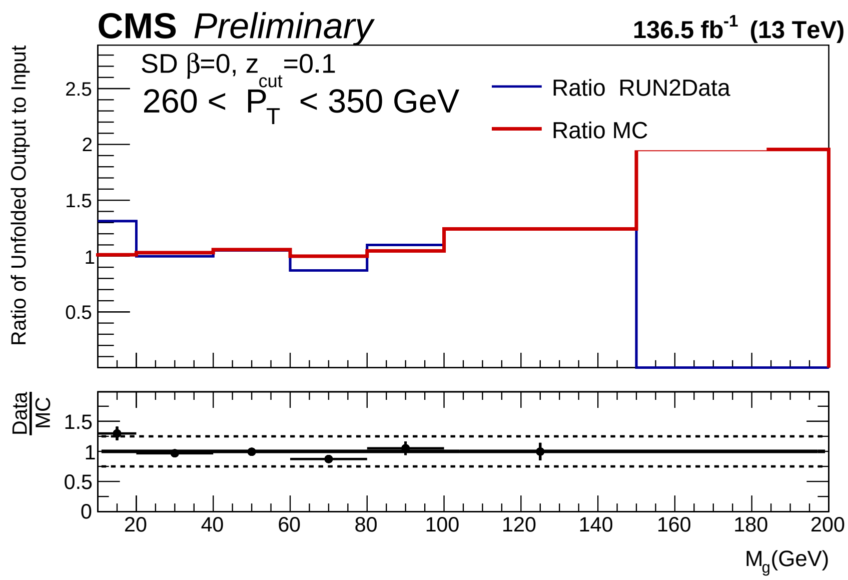

Figure 5.58: Ratio of unfolded over raw data and MC for groomed jets, p_T 260-350 GeV.

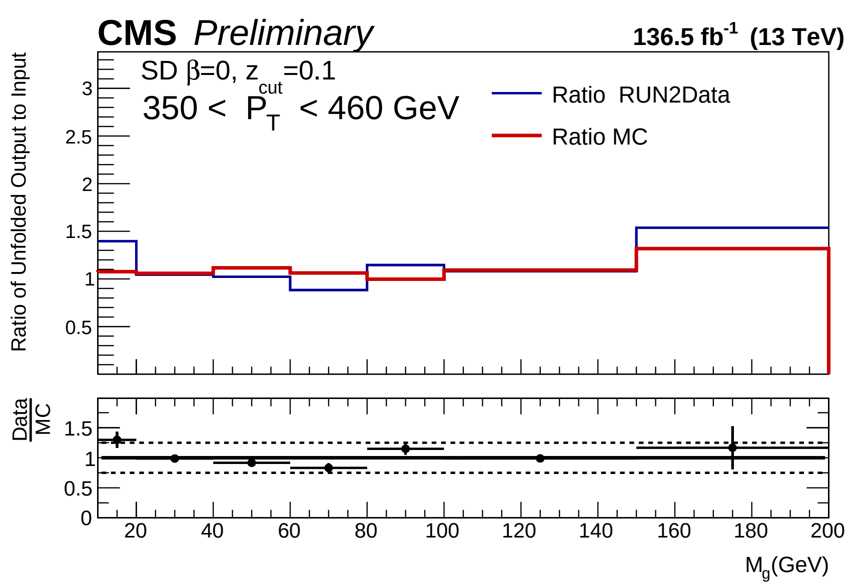

Figure 5.59: Ratio of unfolded over raw data and MC for groomed jets, p_T 350-460 GeV.

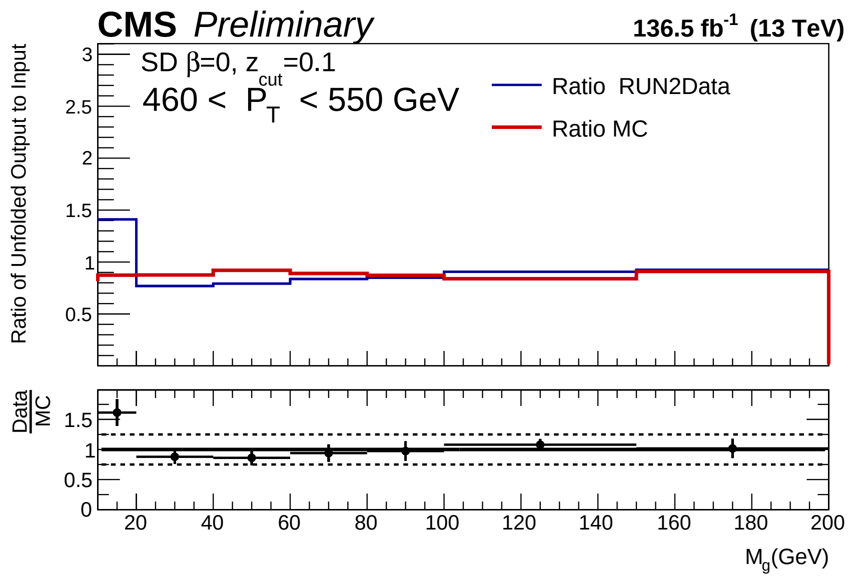

Figure 5.60: Ratio of unfolded over raw data and MC for groomed jets, p_T 460-550 GeV.

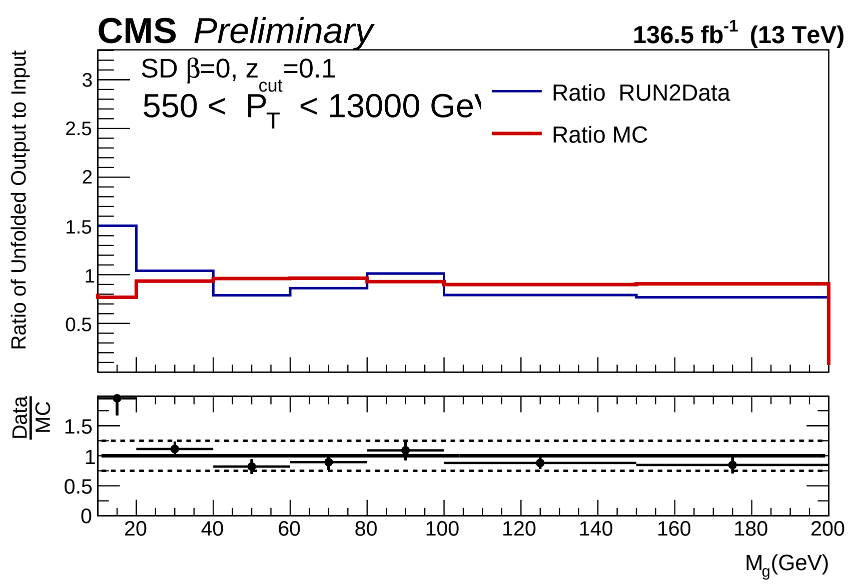

Figure 5.61: Ratio of unfolded over raw data and MC for groomed jets, p_T 550-13000 GeV.

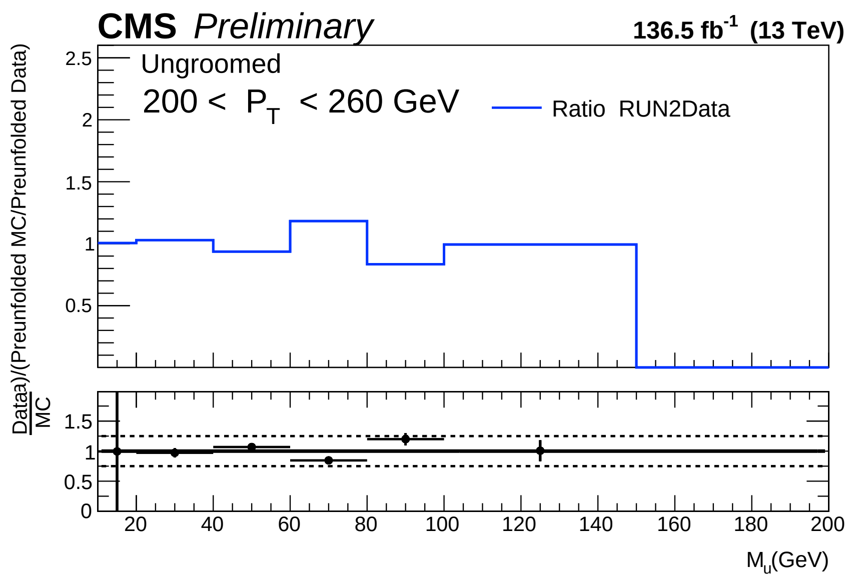

Figure 5.62: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 200-260 GeV.

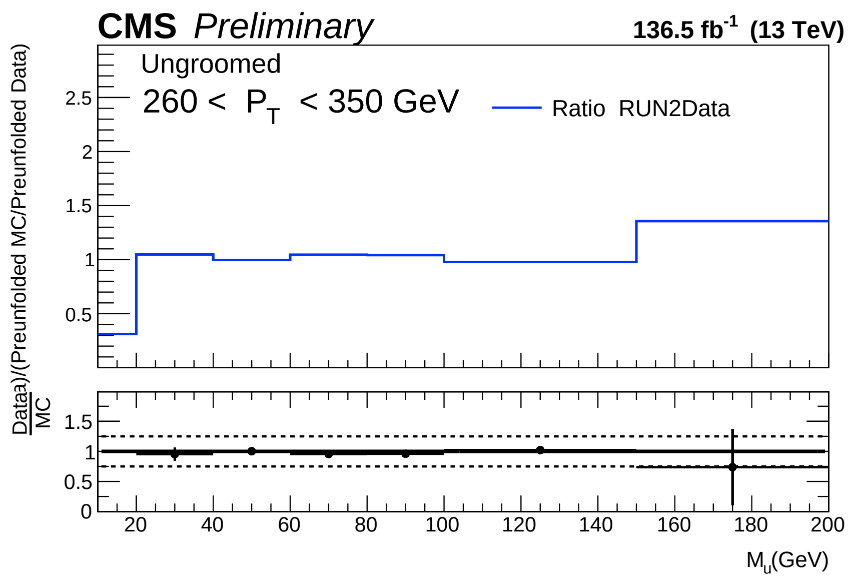

Figure 5.63: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 260-350 GeV.

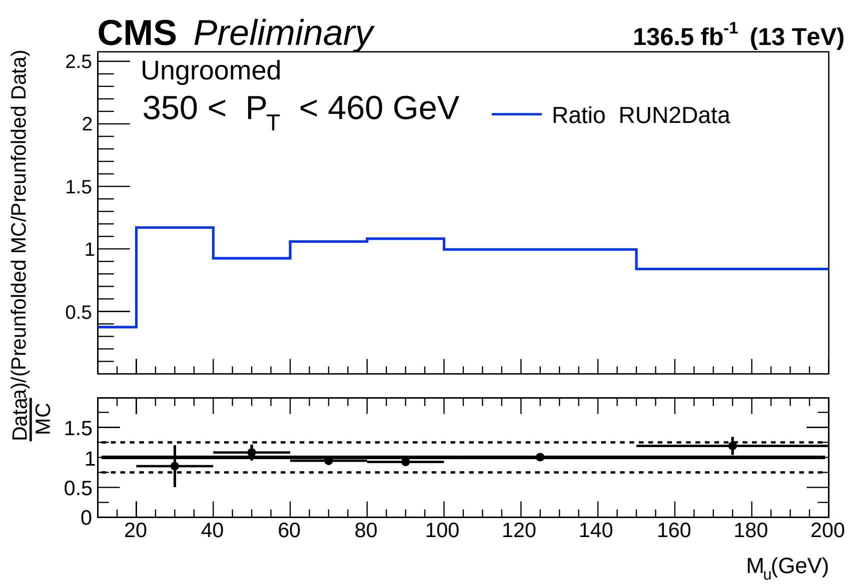

Figure 5.64: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 350-460 GeV.

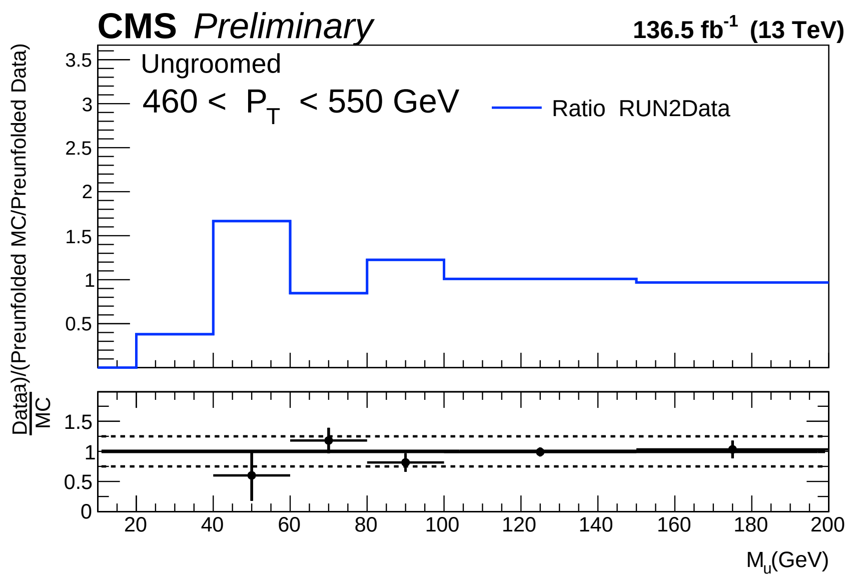

Figure 5.65: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 460-550 GeV.

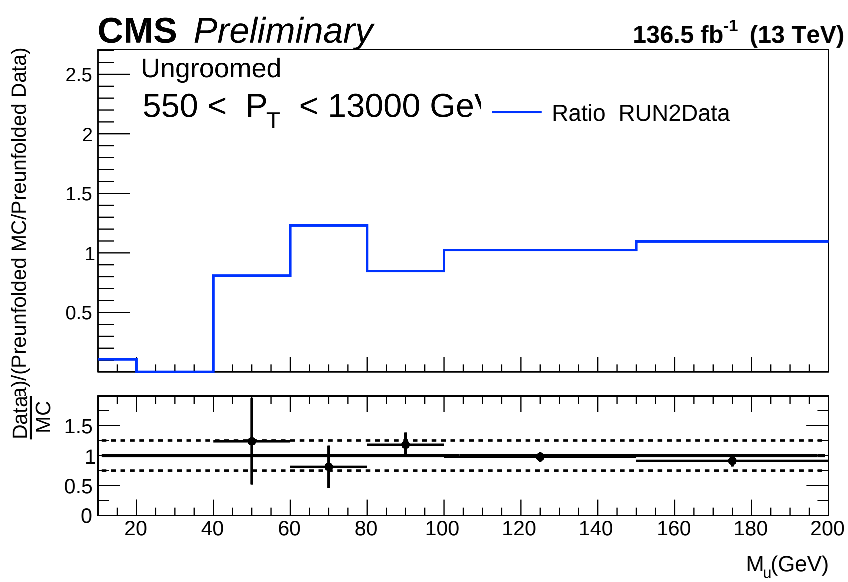

Figure 5.66: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 550-13000 GeV.

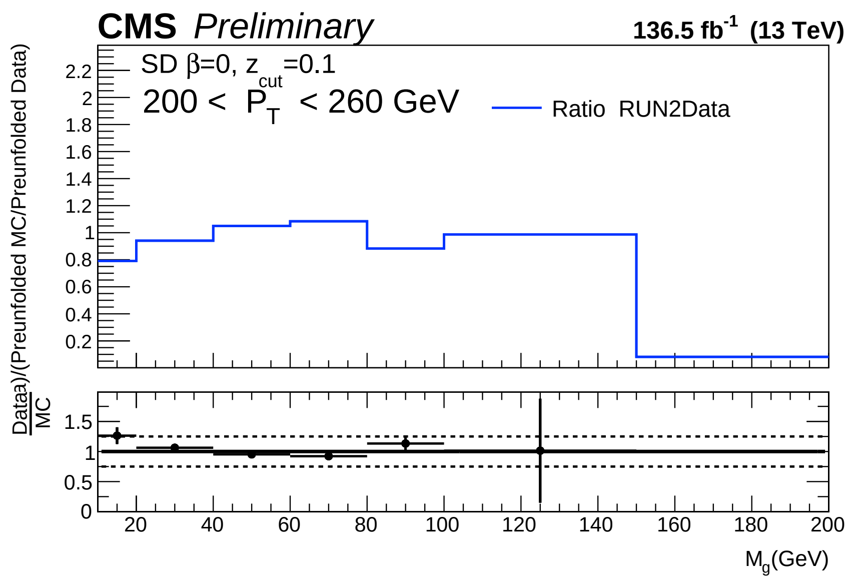

Figure 5.67: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_T 200-260 GeV.

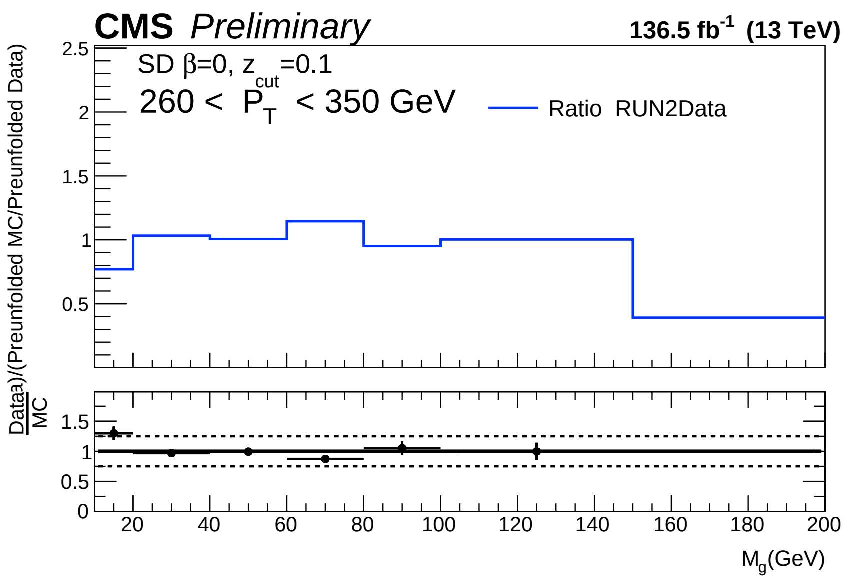

Figure 5.68: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_T 260-350 GeV.

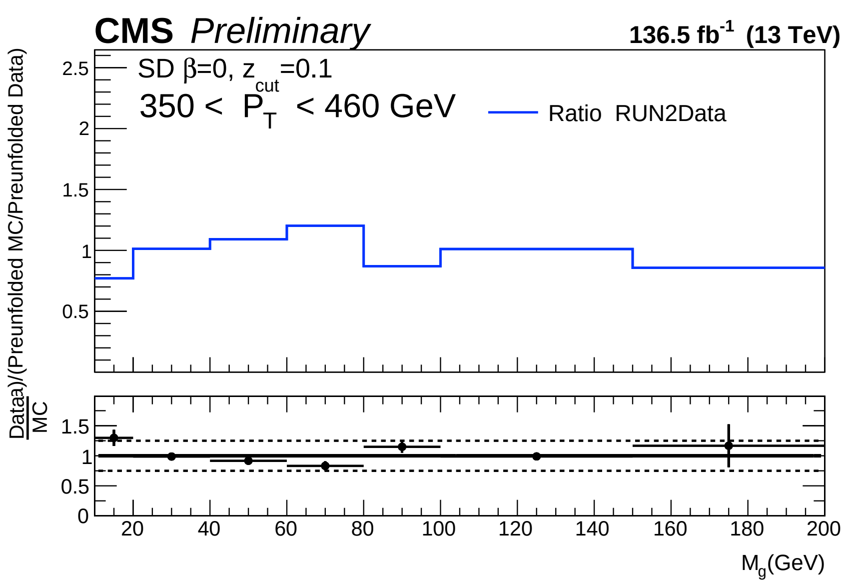

Figure 5.69: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_T 350-460 GeV.

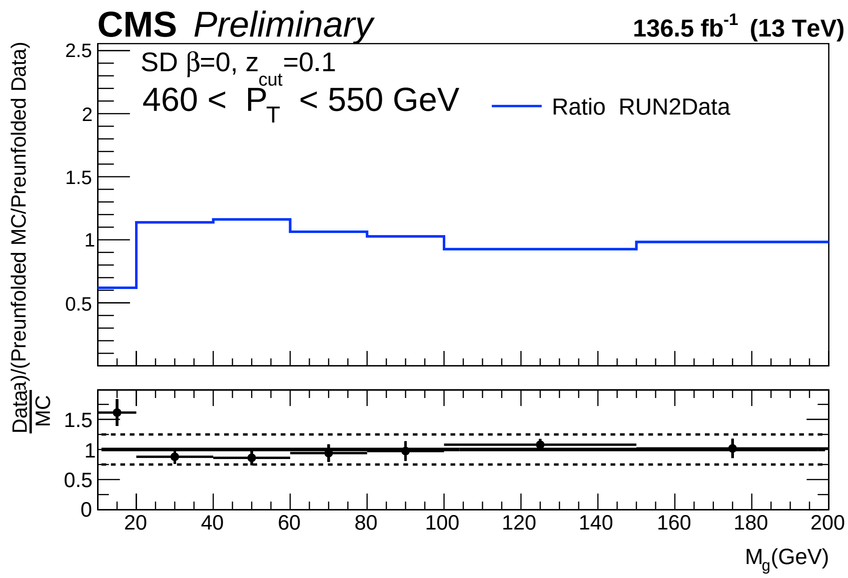

Figure 5.70: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 460-550 GeV.

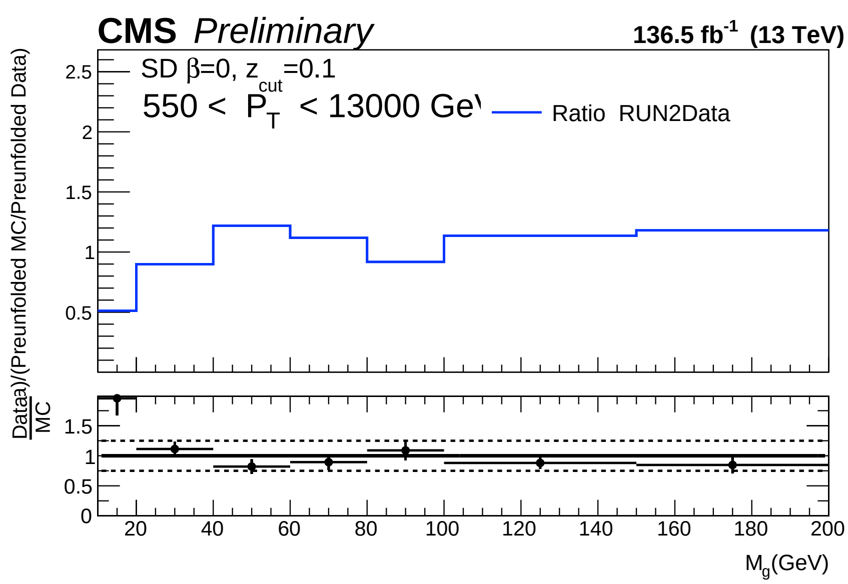

Figure 5.71: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 550-13000 GeV.

5.10 Summary

In conclusion, we have presented a differential jet cross section measured in Z + Jet events in bins of the ungroomed jet p_{T} in conjunction with the ungroomed and groomed jet mass (as well as dimensionless mass) using the “soft drop” (a.k.a. “modified mass drop tagger”) algorithm with 9 different combinations of parameters. The results are presented as the normalized cross section, normalized per reconstructed jet p_{T} bin. Overall leading-order MC simulation agrees reasonably well with the data within our uncertainties. Agreement below the Sudakov peak is slightly worse than above. The application of a grooming algorithm improves the overall precision, with larger improvement at low jet masses. This analysis improves over previous iterations by using various parameter values for the “soft drop” jet grooming algorithm, as well as by including an additional unfolding in both transverse momentum and dimensionless mass, as was done by the ATLAS collaboration [3] .

.1 2016 data results

This Appendix shows the distributions from the ”Detector Response” through the ”Results” sections of the main analysis note with only the 2016 data rather than the full Run 2 statistics seen in the main body of the note.

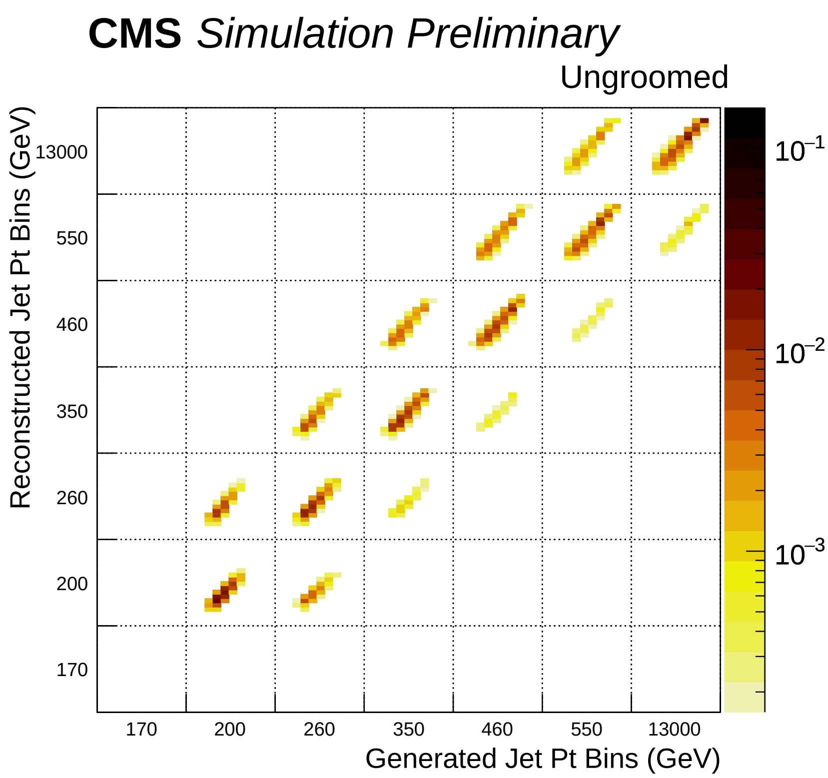

Figure 72: Two-dimensional response matrix for ungroomed jets. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV). While the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

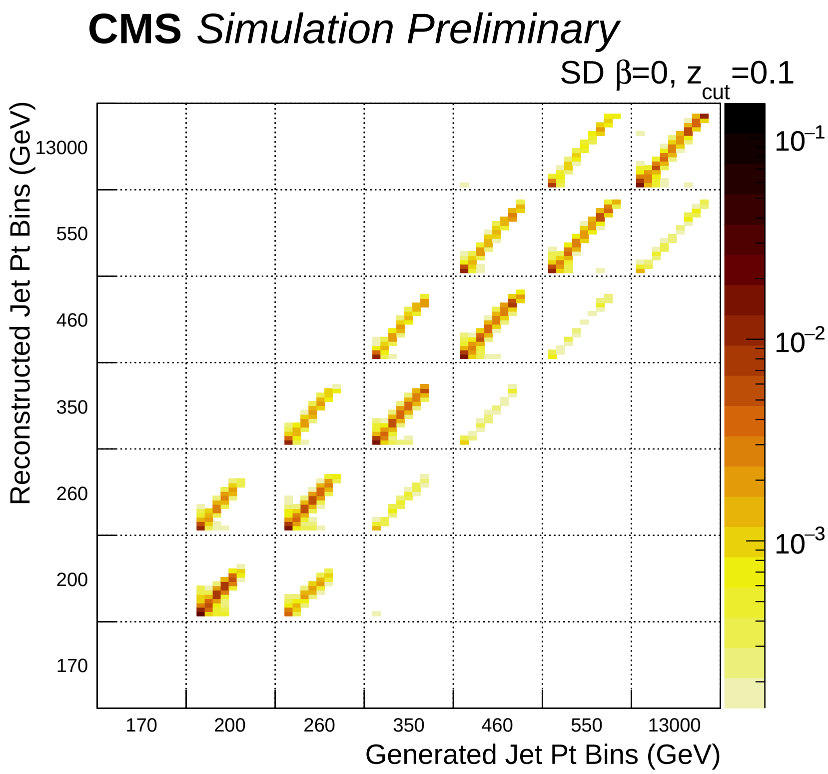

Figure 73: Two-dimensional response matrix for groomed jets $\beta = 0$, $z_{cut} = 0.1$. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV).while the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

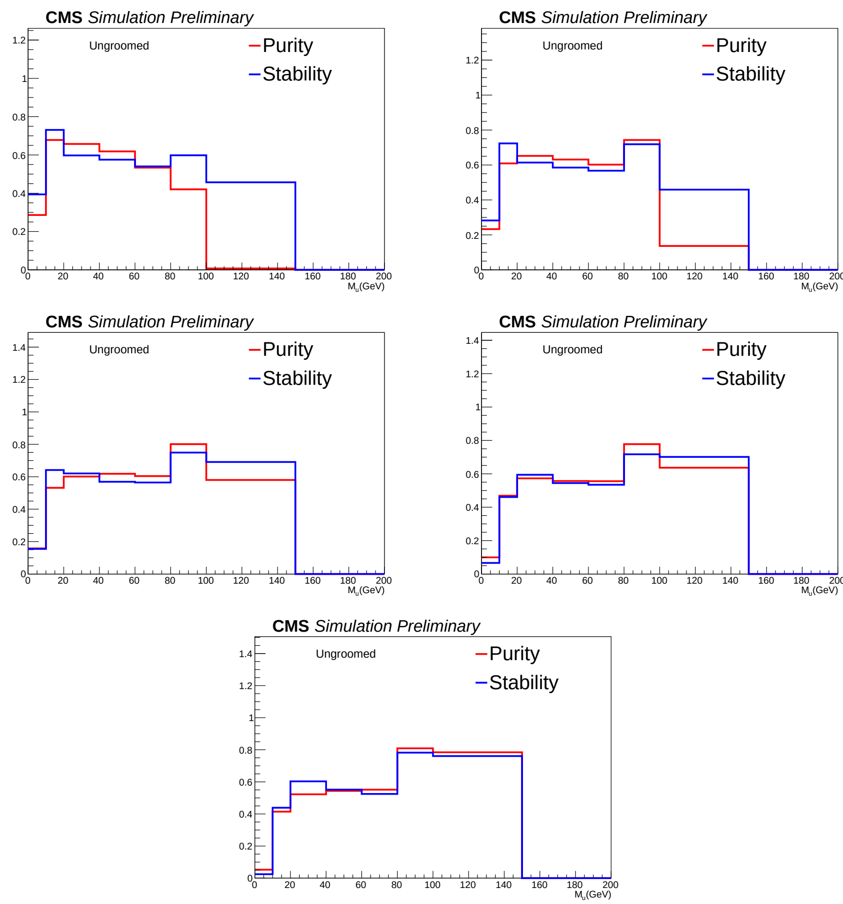

Figure 74: Purity and stability for ungroomed jets in various p_{T} bins.

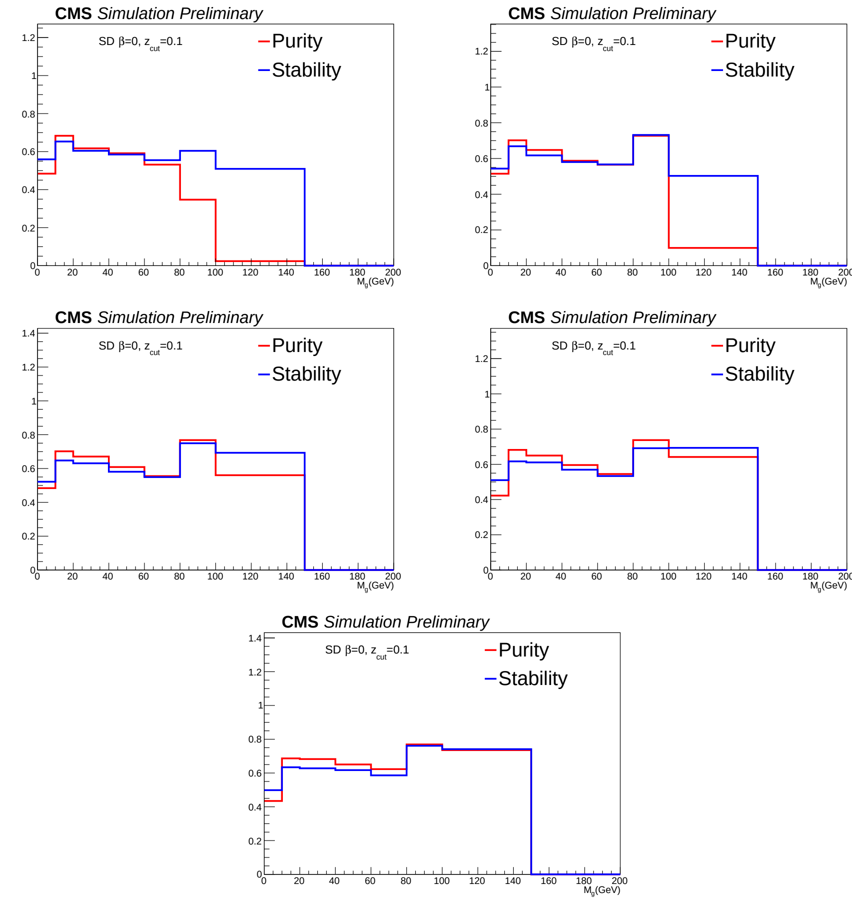

Figure 75: Purity and stability for groomed jets, where the soft-drop criterion was applied with $\beta = 0$ and $z_{cut} = 0.1$, in various p_{T} bins.

Figure-76 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_T bin for the normalized cross section.

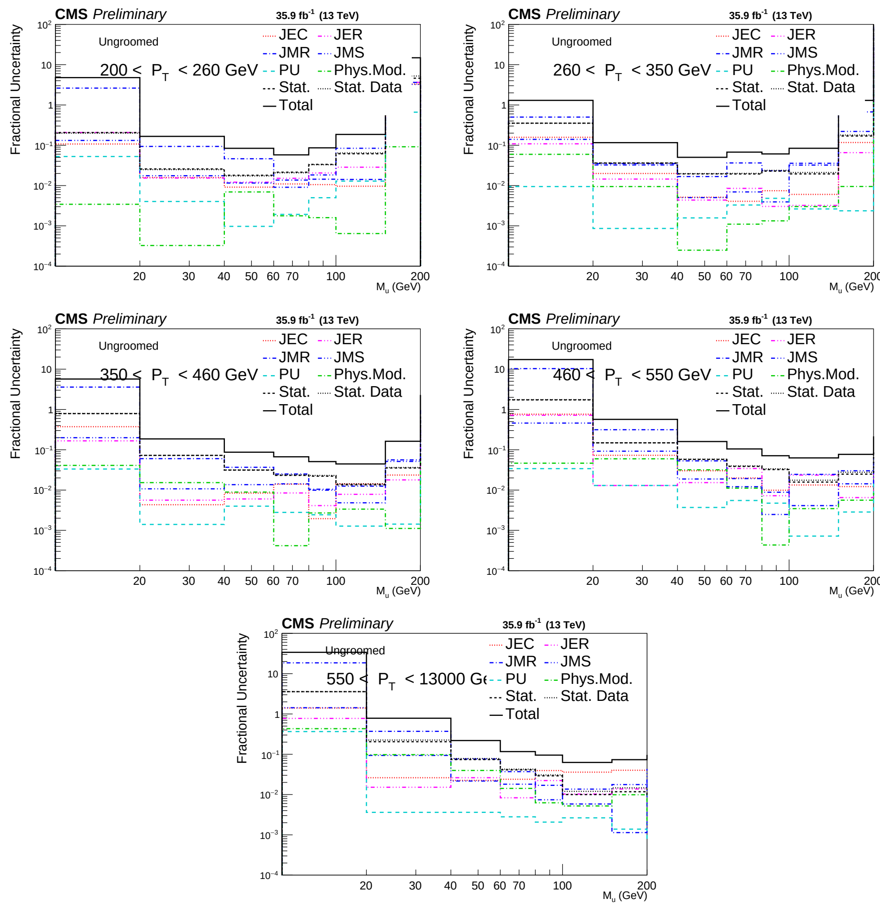

Figure 76: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the raw jet mass.

Figure-78 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_T bin for the normalized cross section.

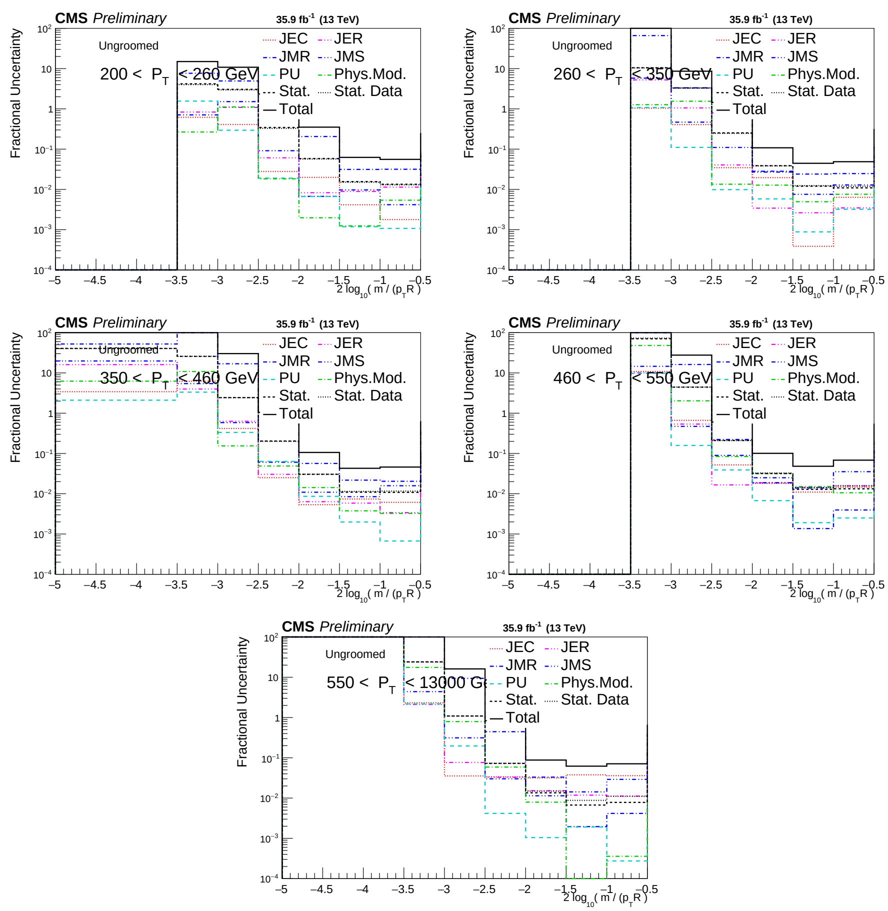

Figure 77: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_{T} bins for the raw dimensionless jet mass.

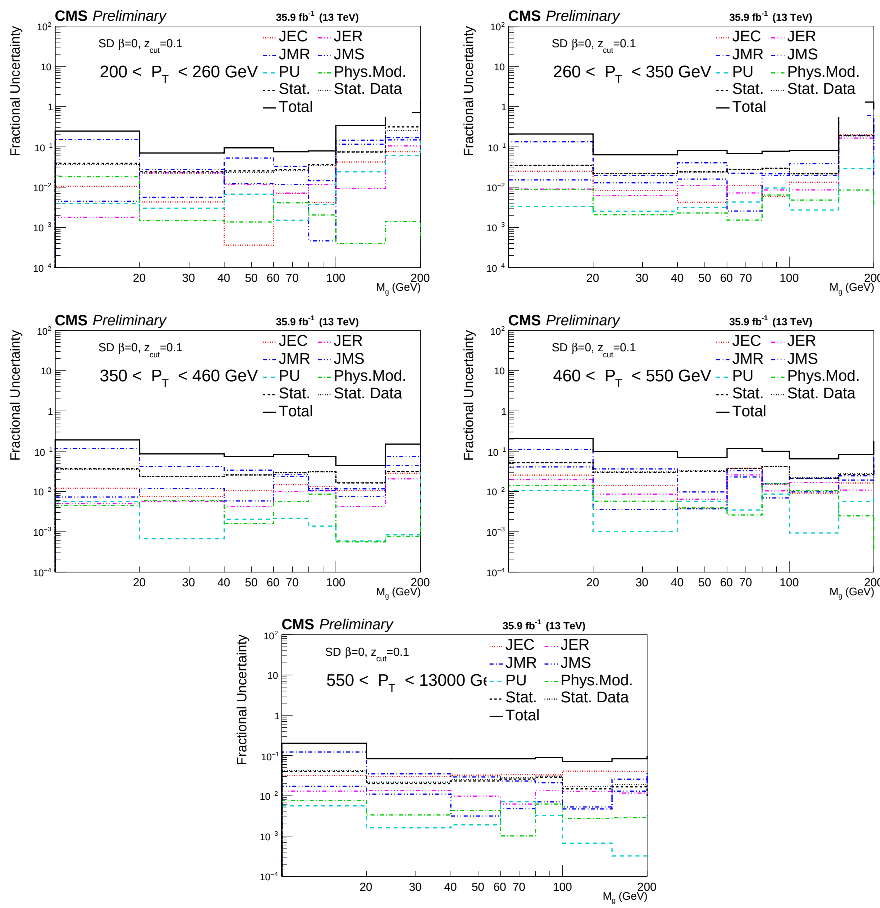

Figure 78: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the jet mass after grooming with default Soft-Drop.

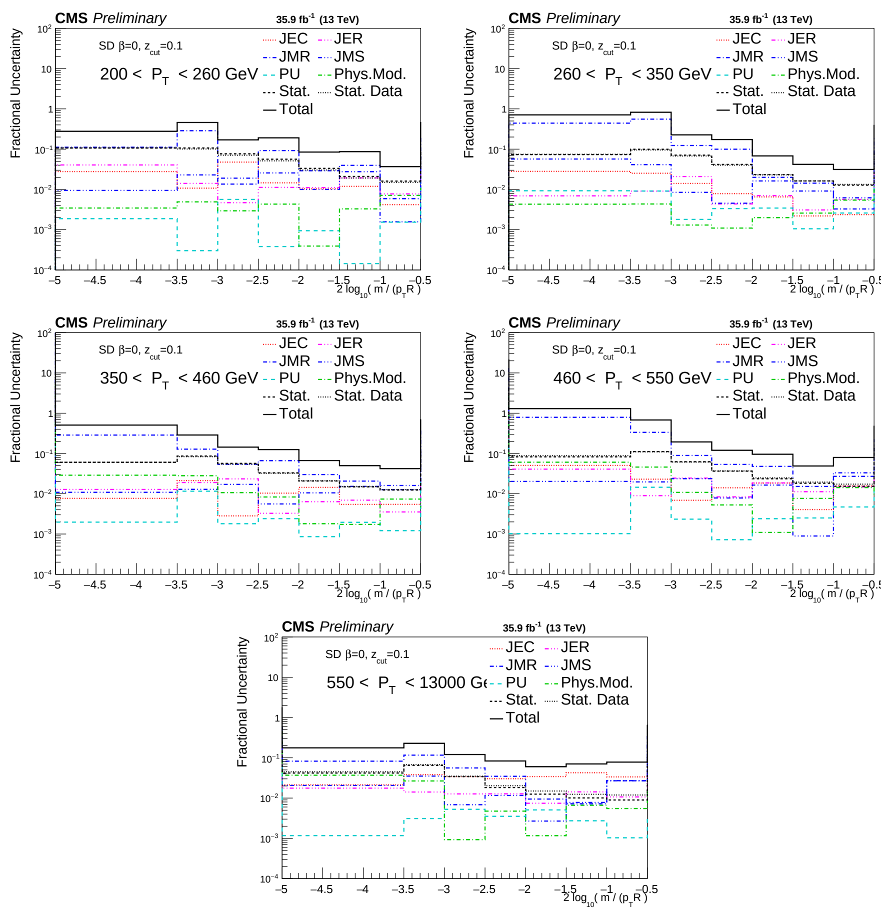

Figure 79: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the jet mass after grooming with default Soft-Drop.

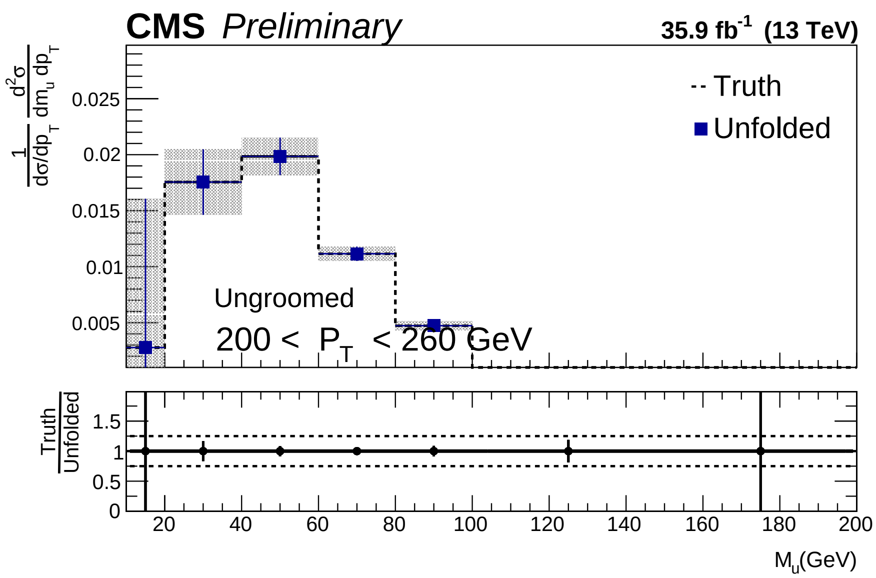

Figure 80: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

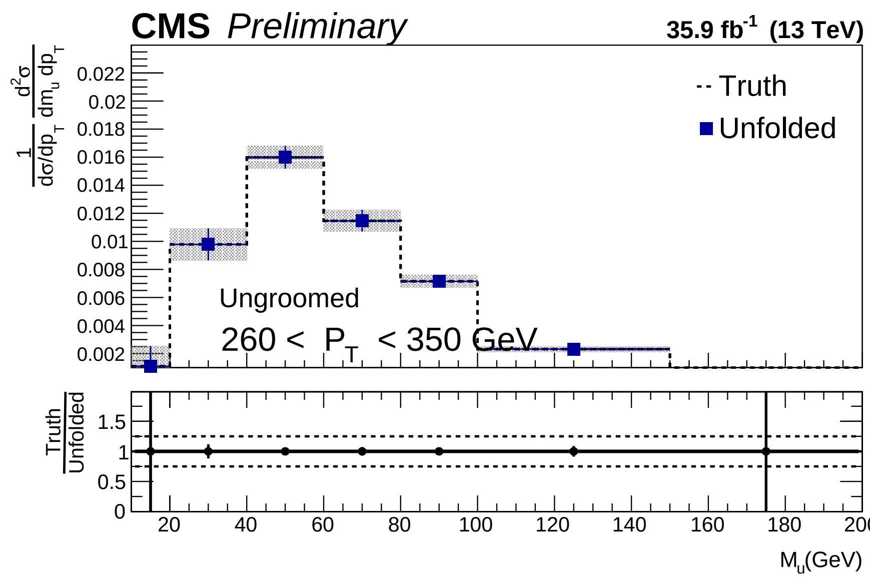

Figure 81: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

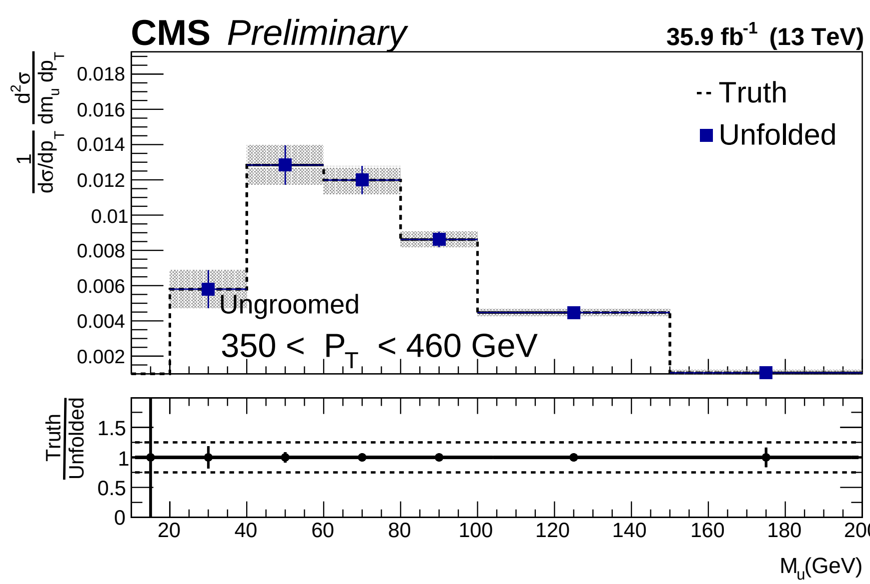

Figure 82: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

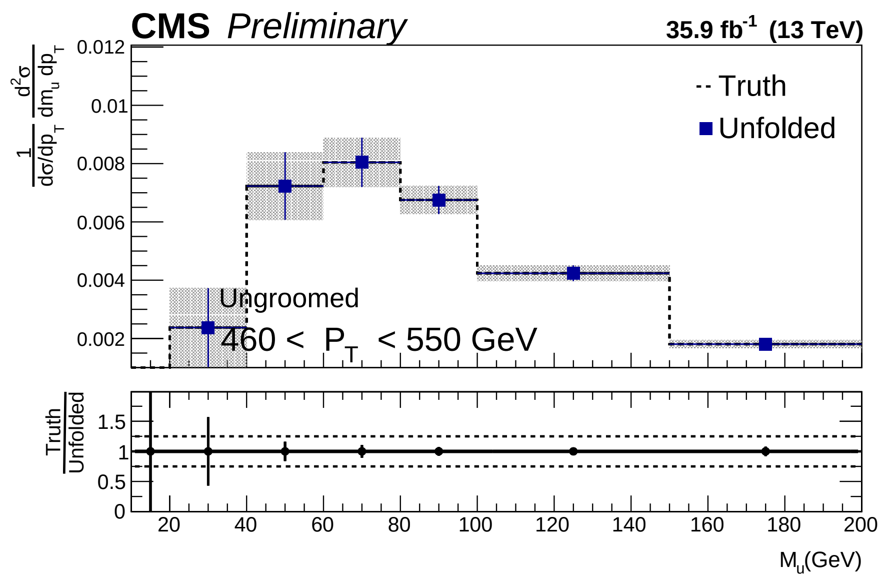

Figure 83: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

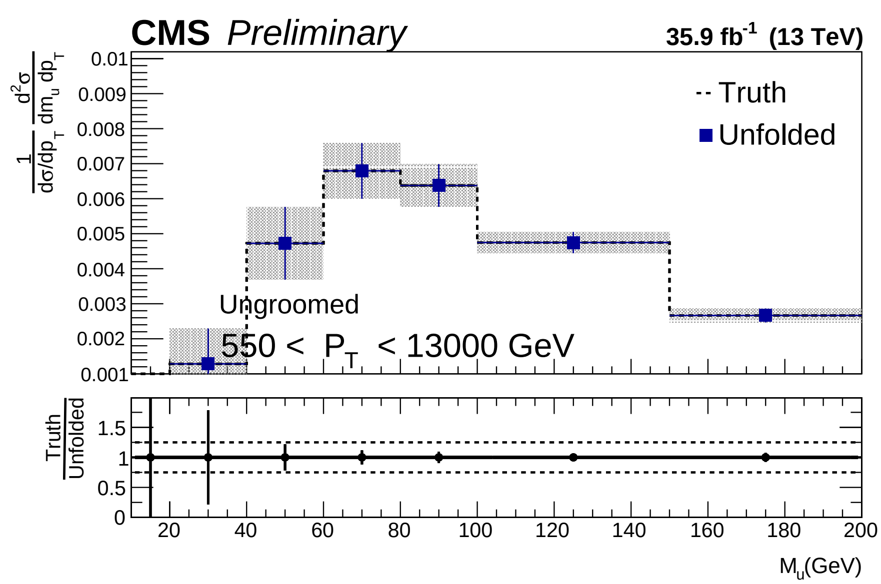

Figure 84: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

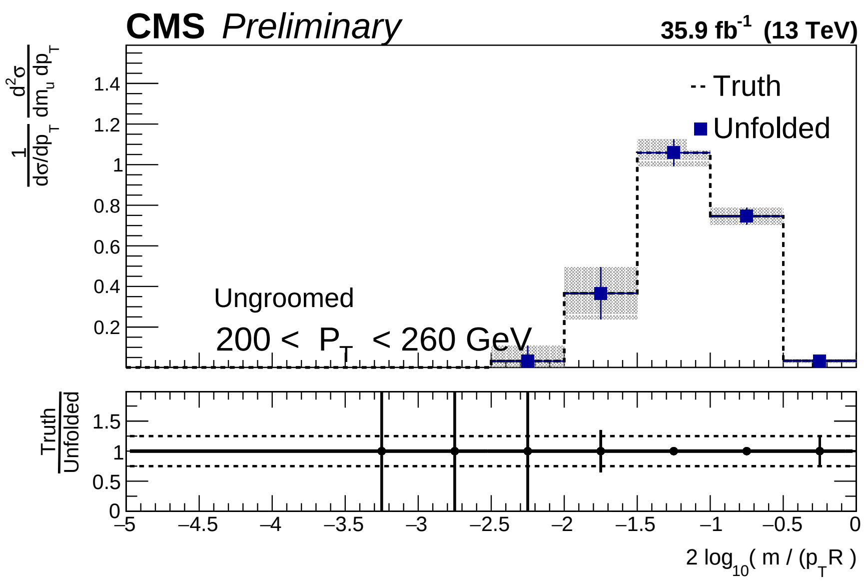

Figure 85: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

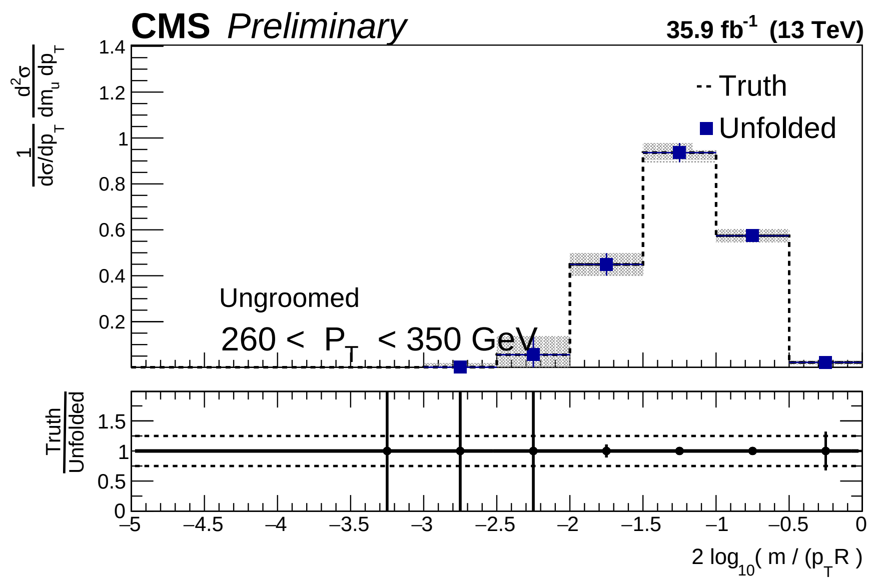

Figure 86: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 260-350 GeV.

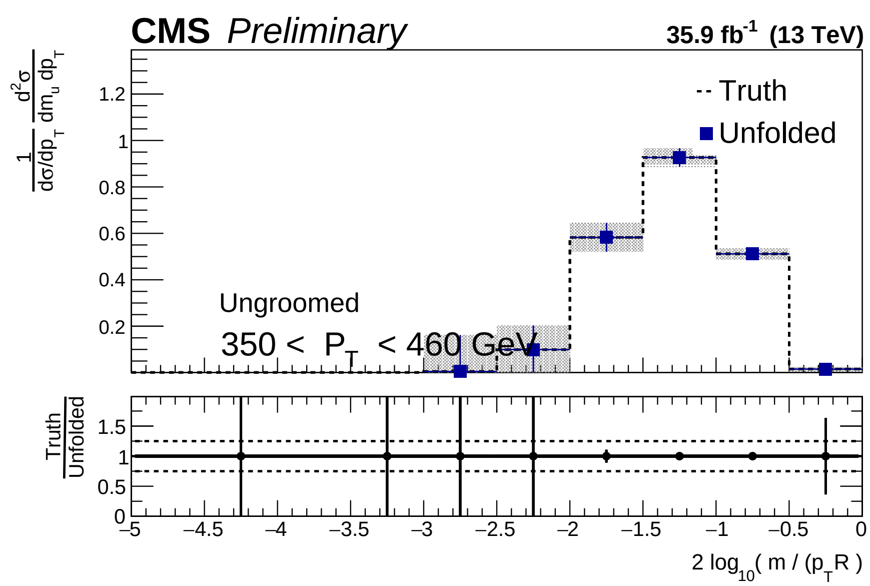

Figure 87: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 350-460 GeV.

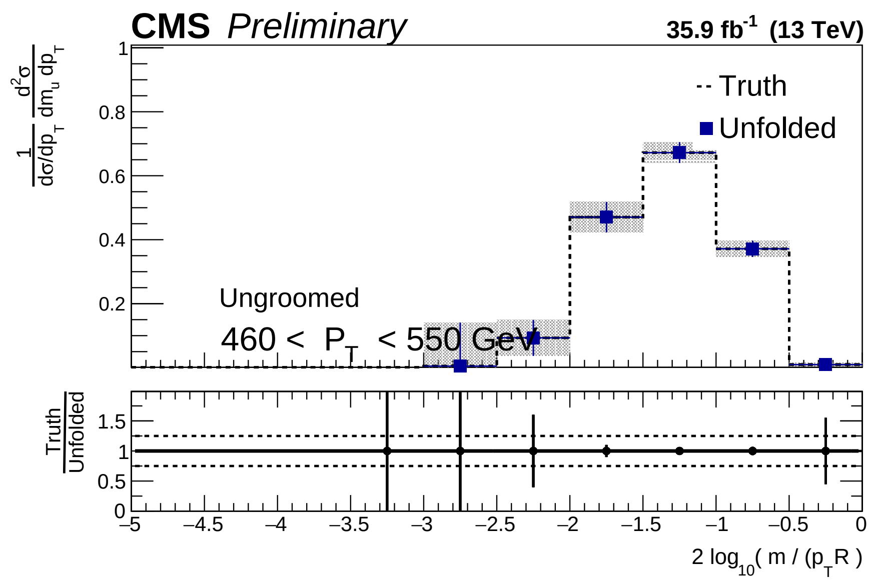

Figure 88: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 460-550 GeV.

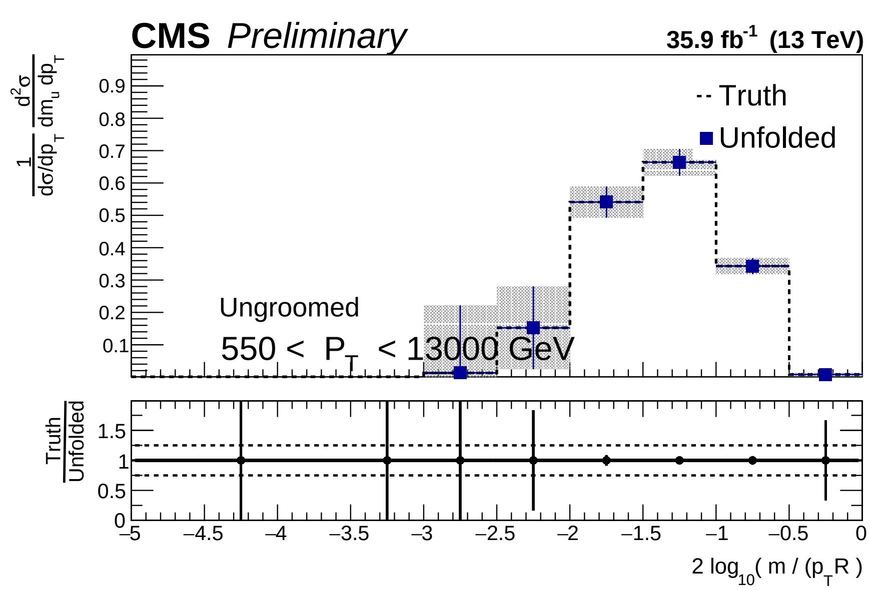

Figure 89: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 550-13000 GeV.

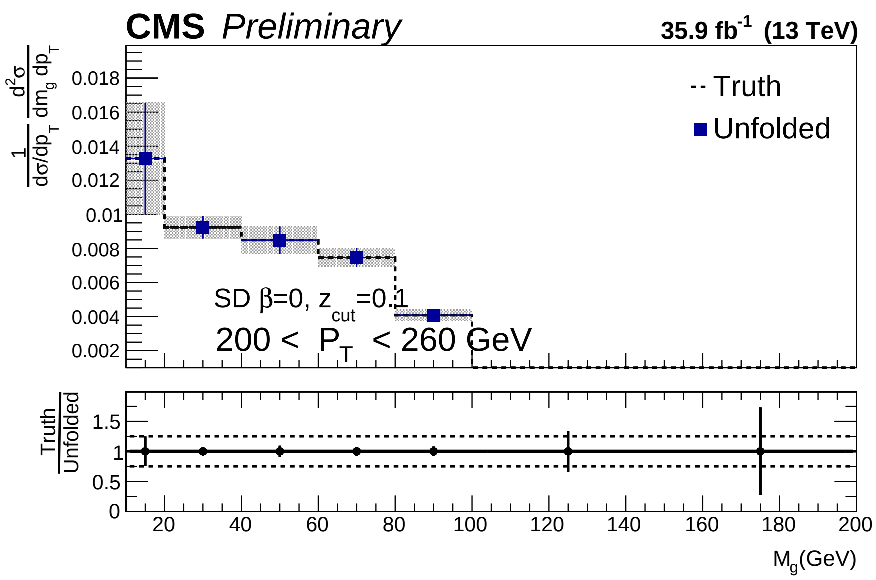

Figure 90: Closure test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

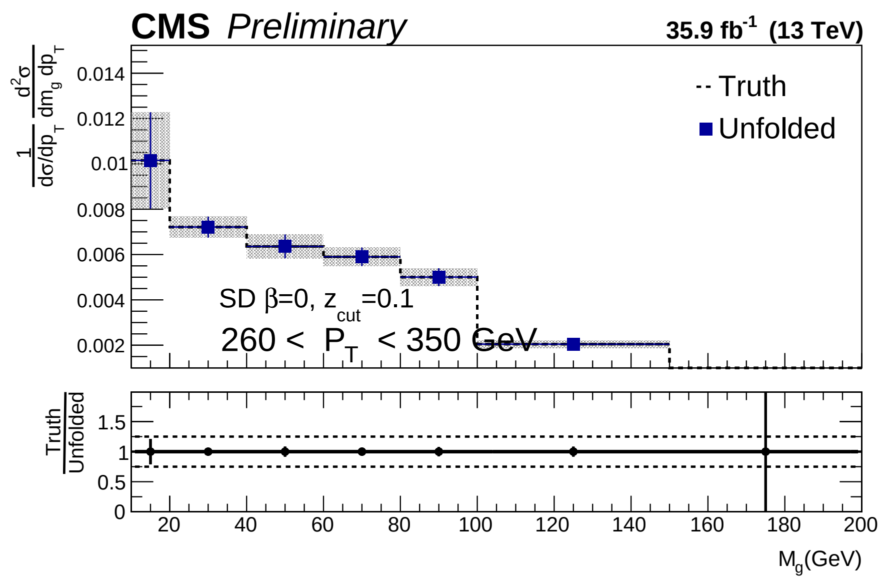

Figure 91: Closure test of groomed reconstructed Monte Carlo, p_T 260-350 GeV.

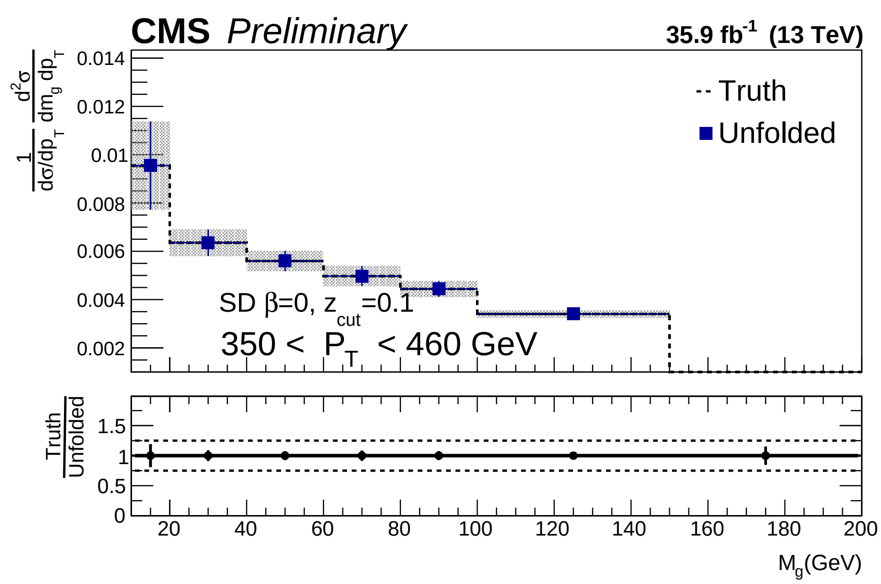

Figure 92: Closure test of groomed reconstructed Monte Carlo, p_T 350-460 GeV.

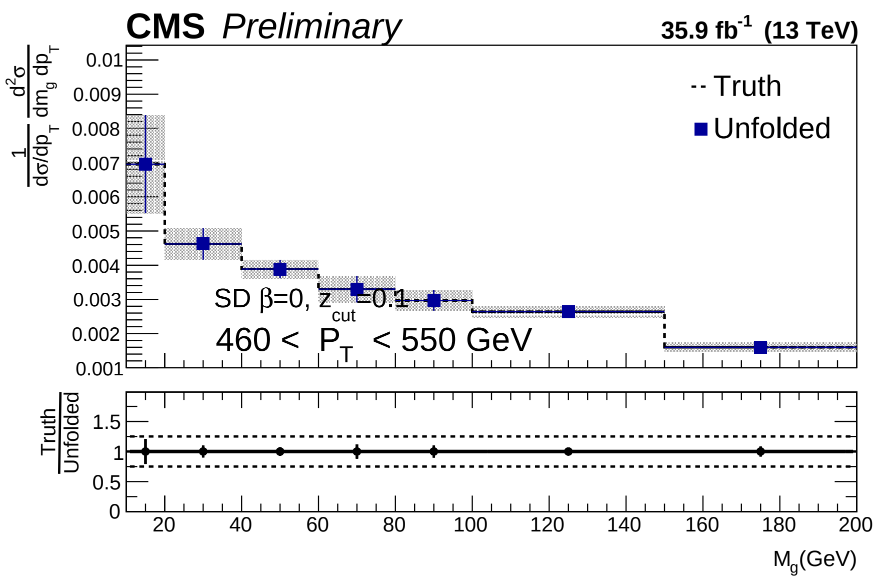

Figure 93: Closure test of groomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

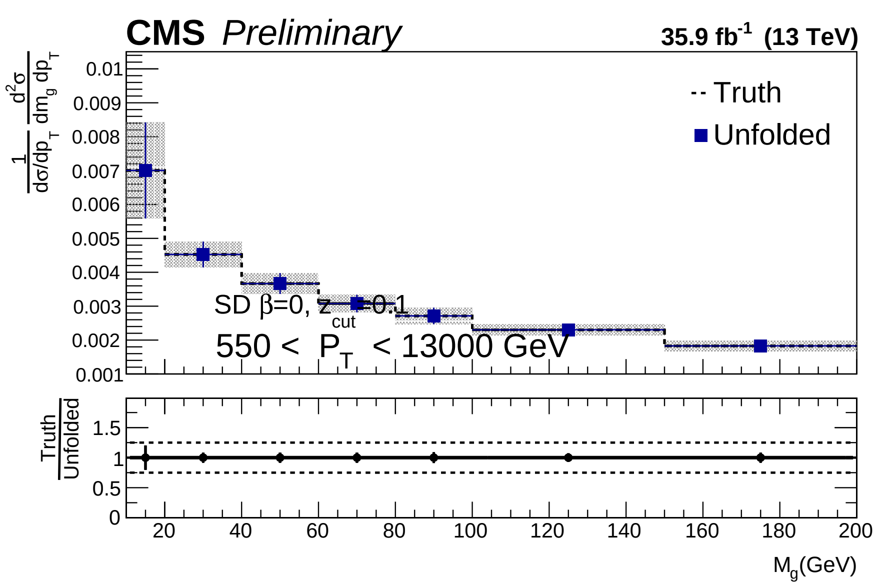

Figure 94: Closure test of groomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

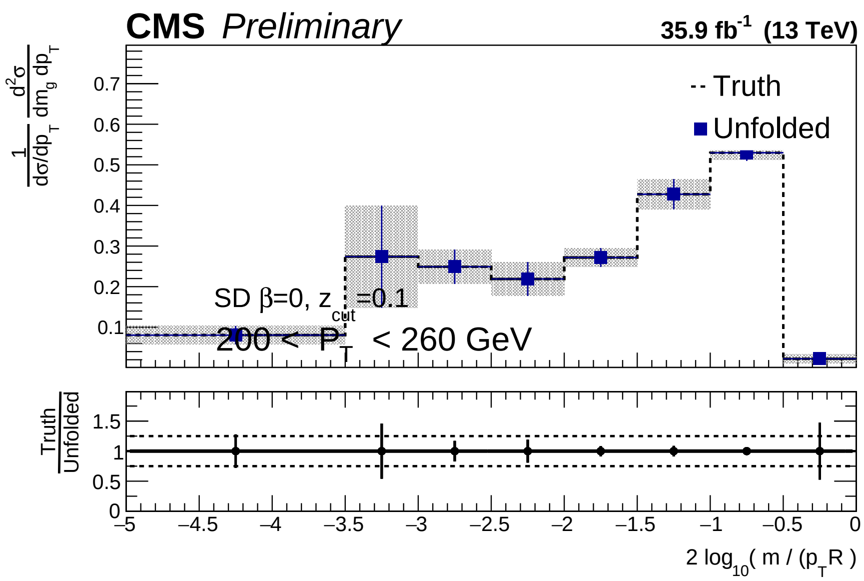

Figure 95: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_T 200-260 GeV.

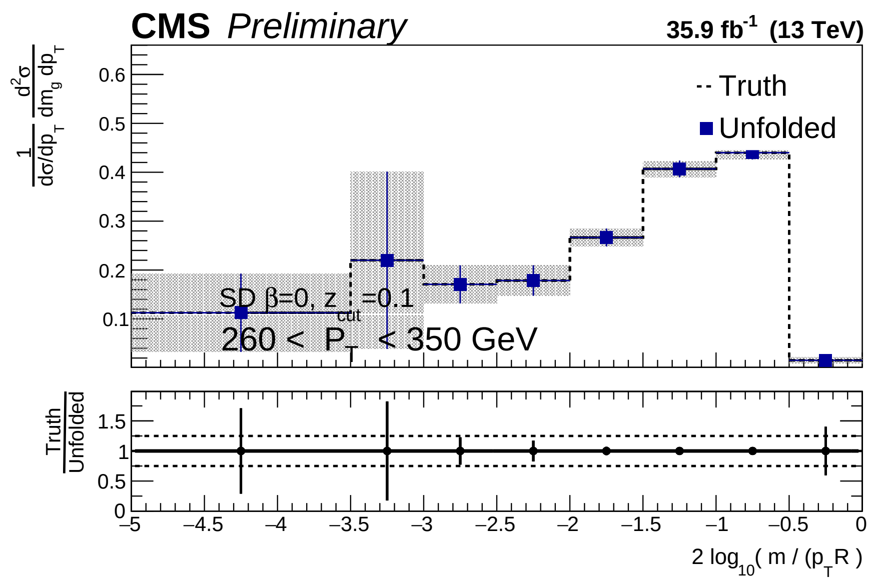

Figure 96: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

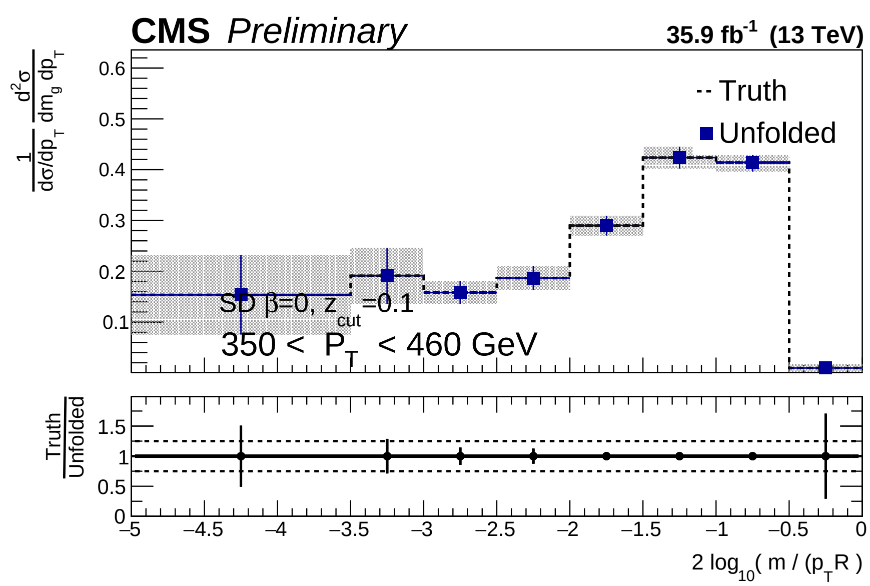

Figure 97: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

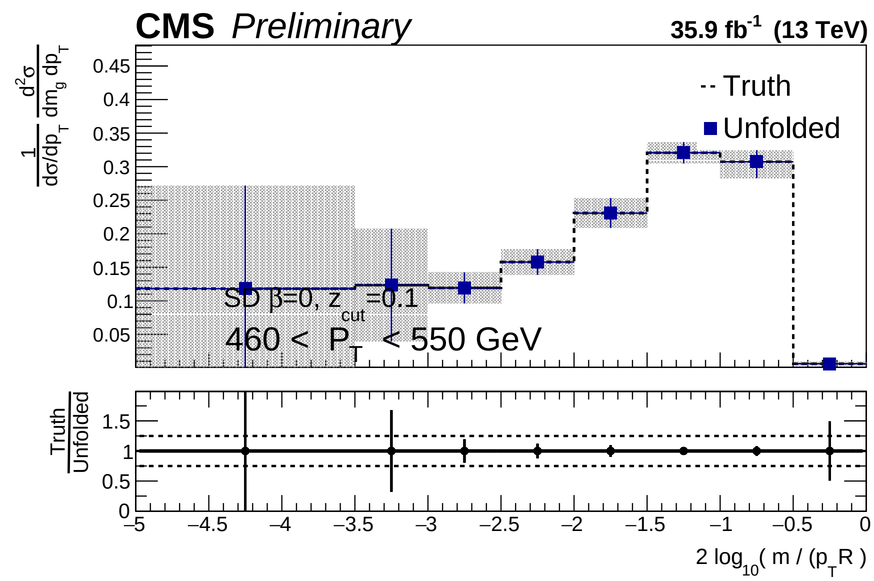

Figure 98: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_T 460-550 GeV.

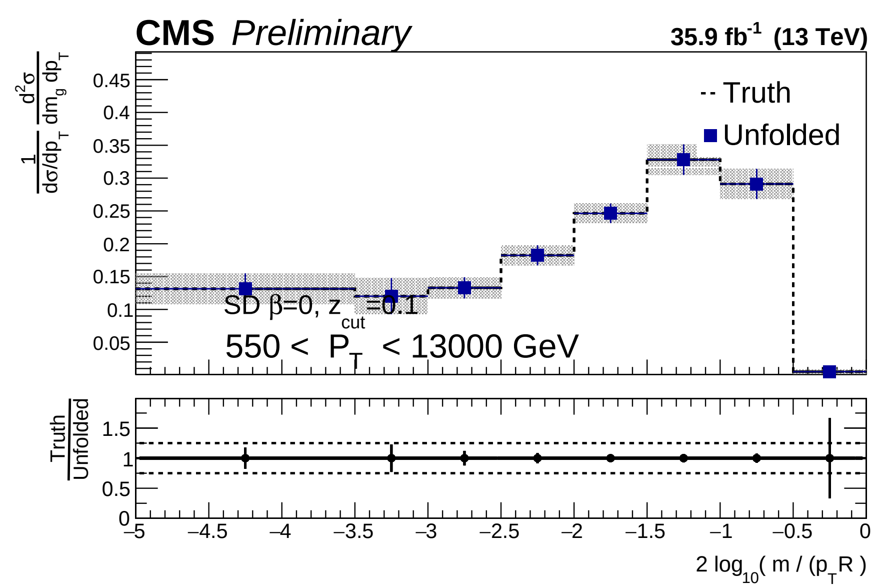

Figure 99: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_T 550-13000 GeV.

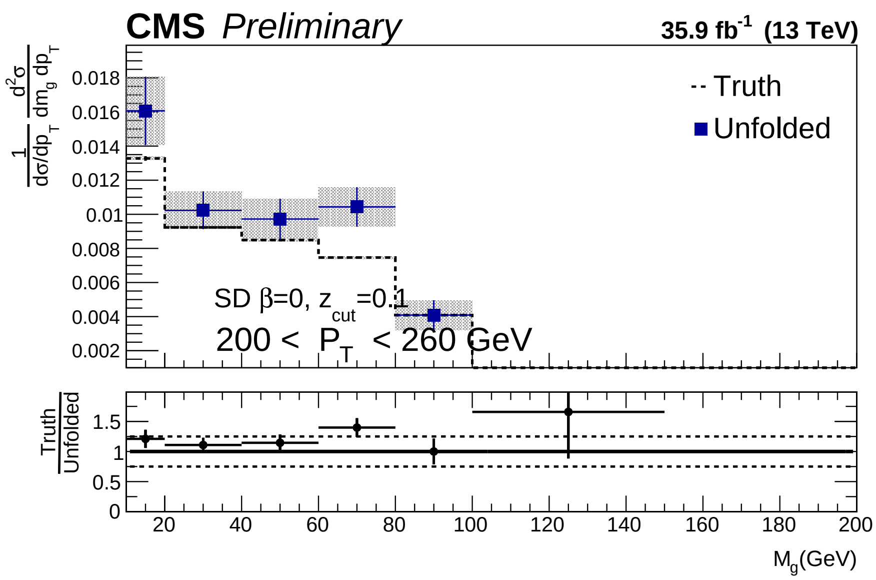

Figure 100: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

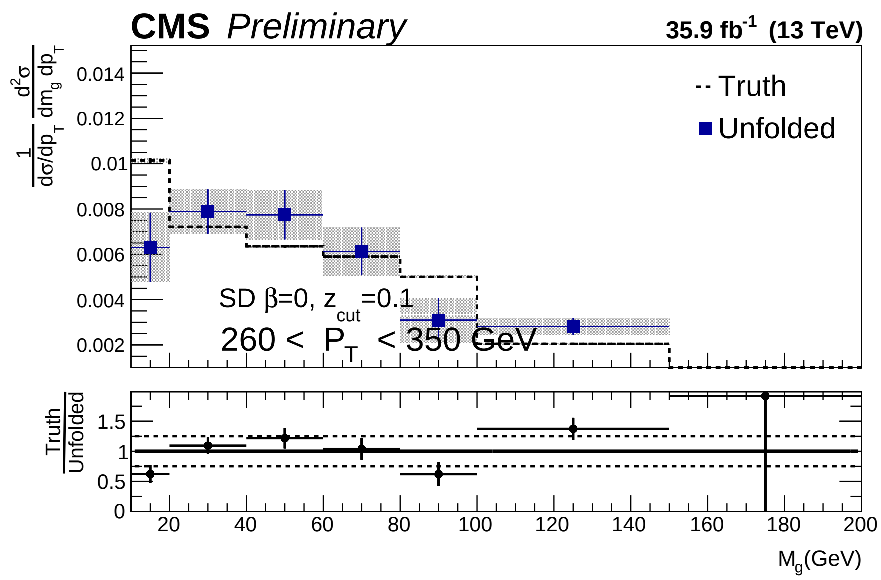

Figure 101: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

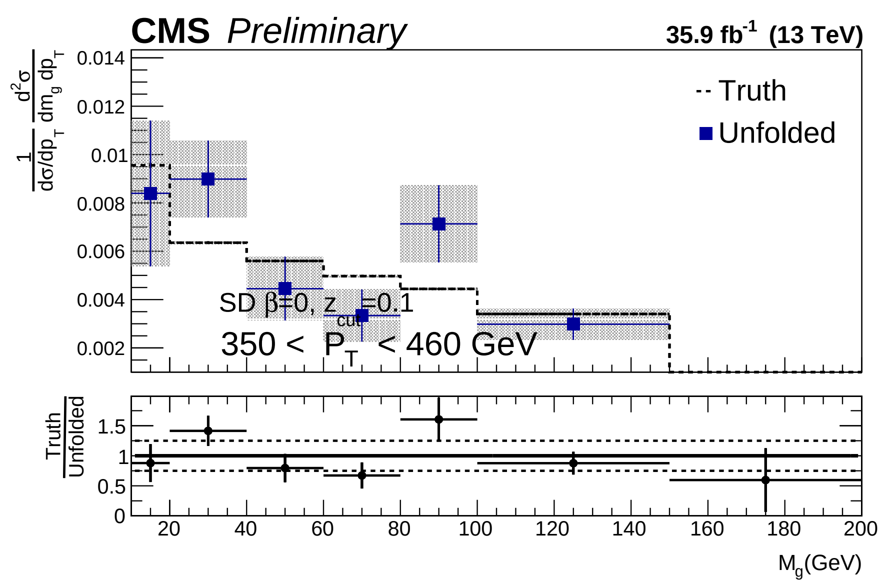

Figure 102: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

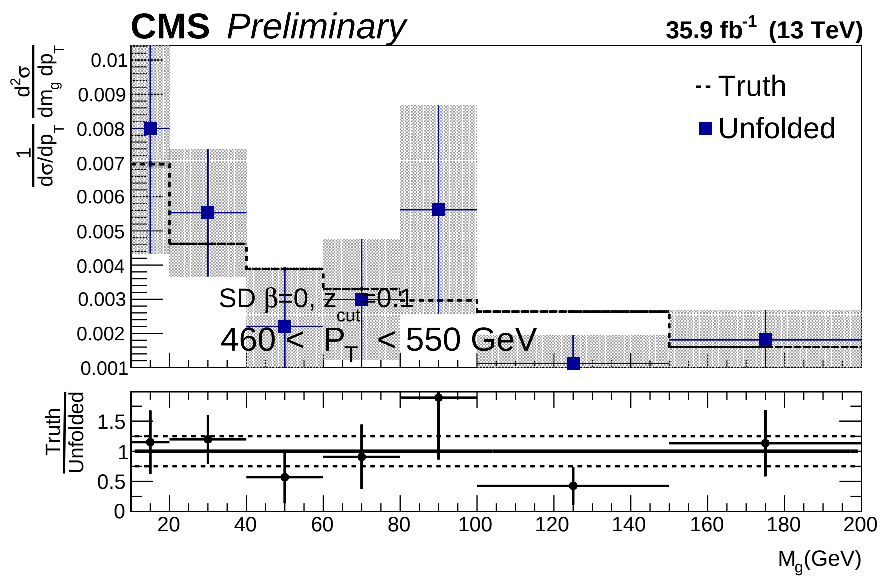

Figure 103: Physical bias test of groomed reconstructed Monte Carlo, p_T 460-550 GeV.

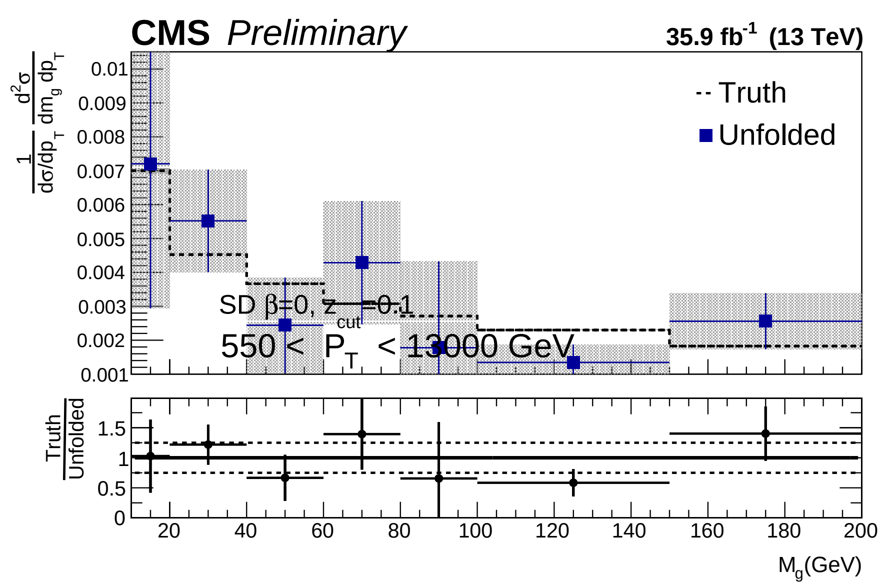

Figure 104: Physical bias test of groomed reconstructed Monte Carlo, p_T 550-13000 GeV.

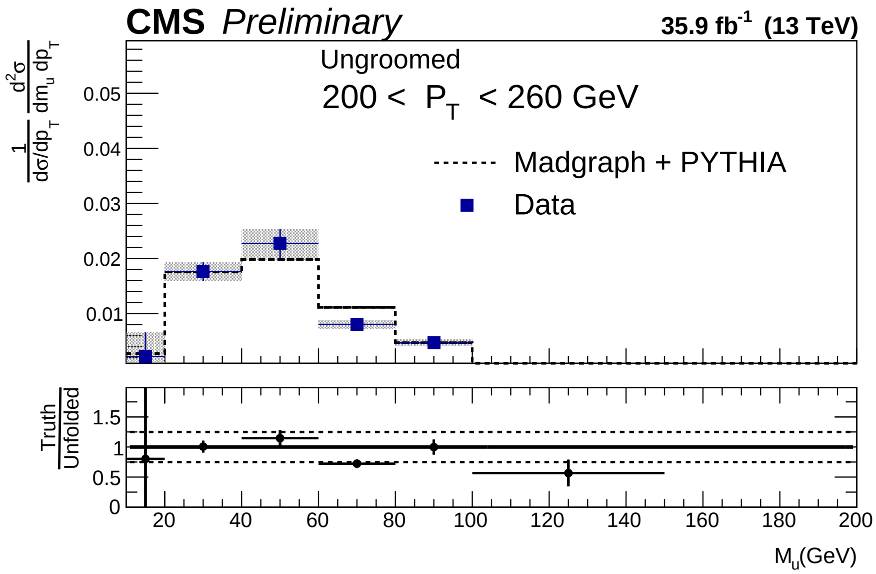

Figure 105: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 200-260 GeV.

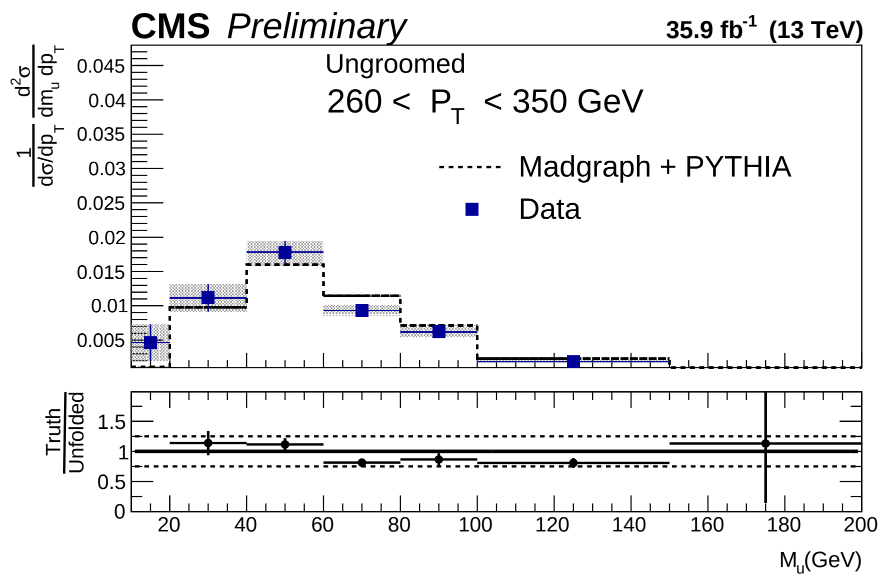

Figure 106: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 260-350 GeV.

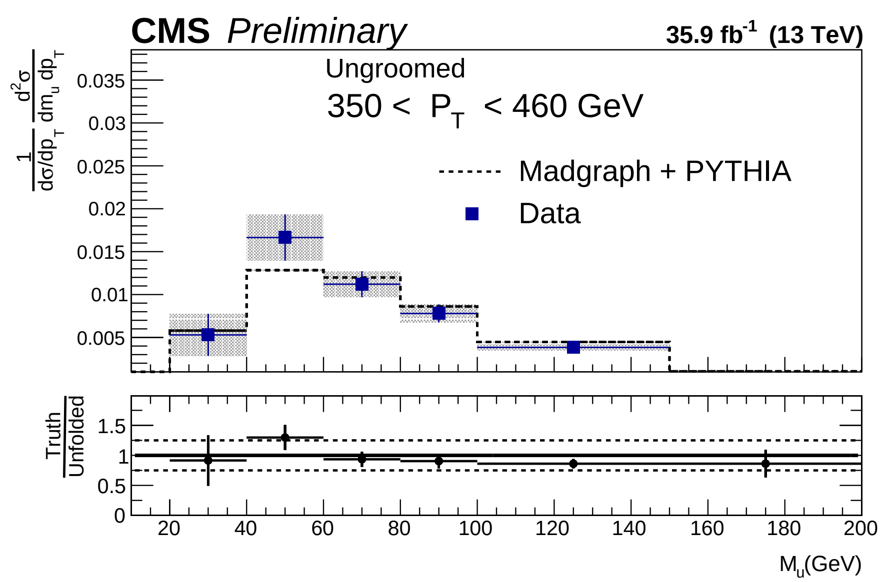

Figure 107: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 350-460 GeV.

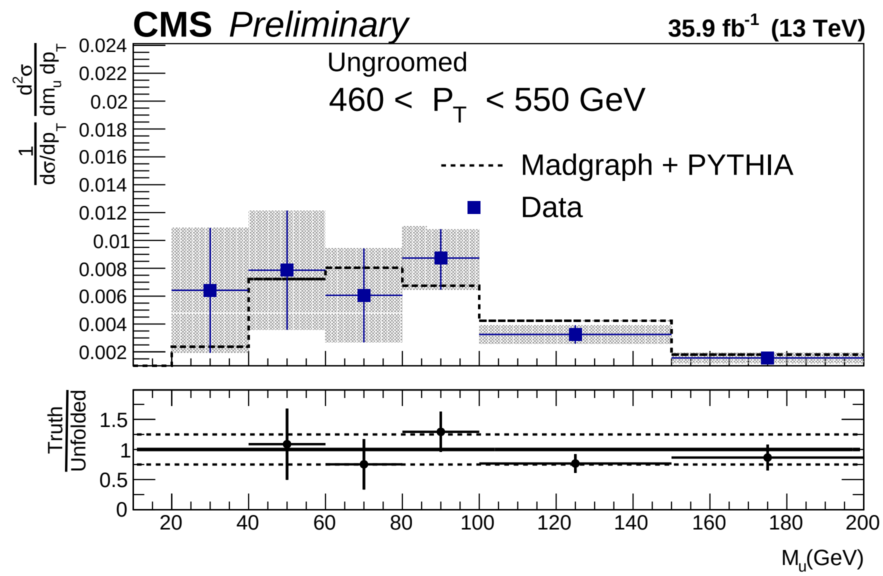

Figure 108: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 460-550 GeV.

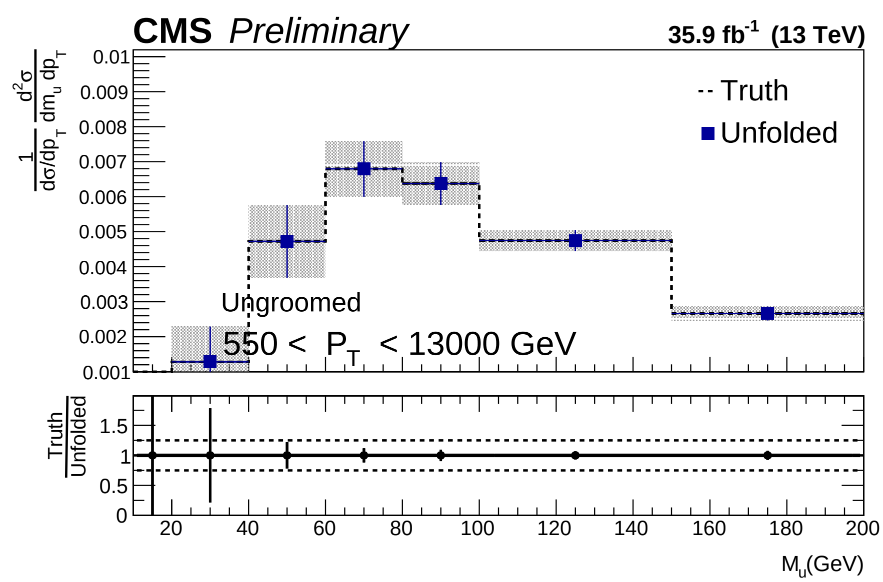

Figure 109: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 550-13000 GeV.

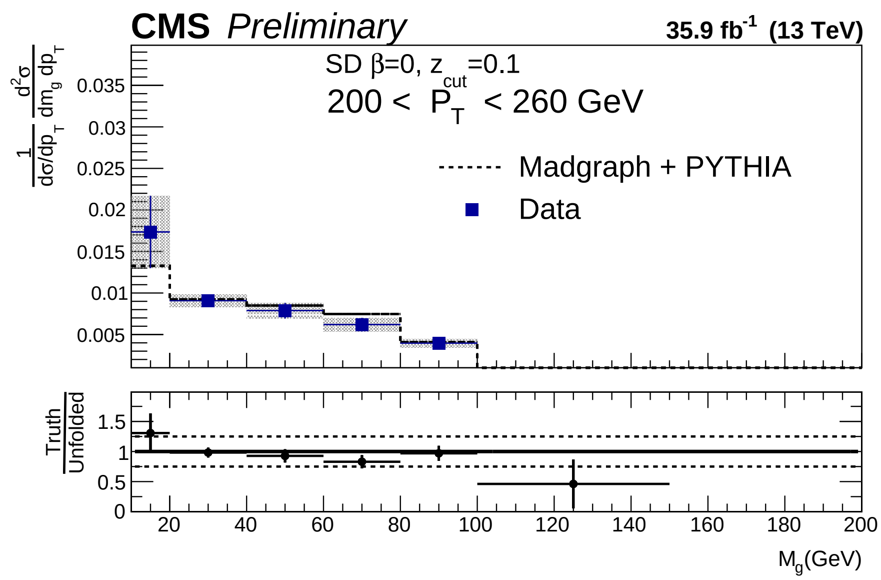

Figure 110: Normalized cross section results with respect to jet mass for groomed jets, p_T 200-260 GeV.

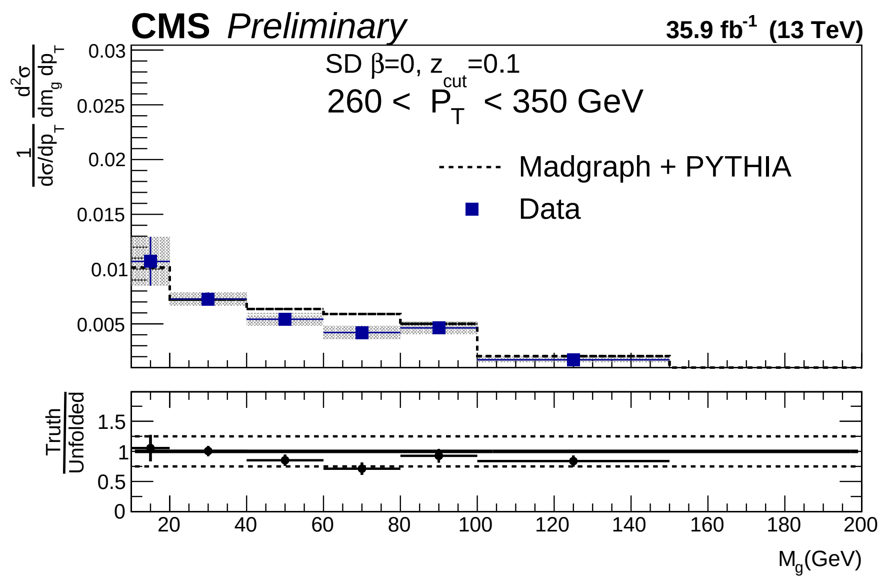

Figure 111: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 260-350 GeV.

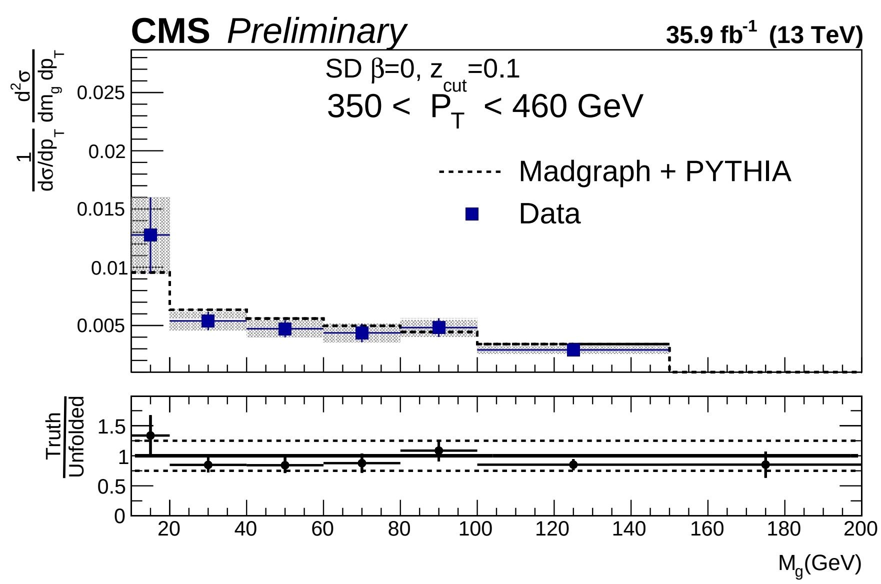

Figure 112: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 350-460 GeV.

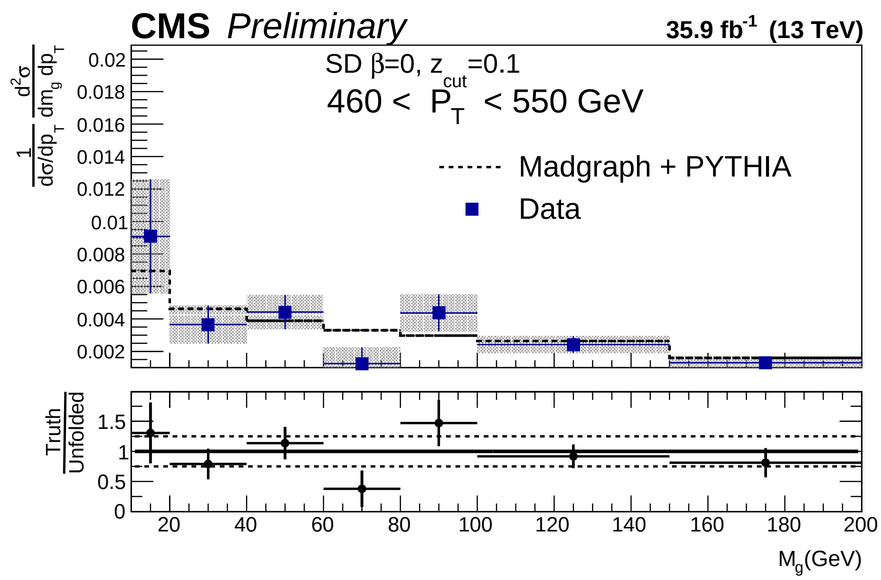

Figure 113: Normalized cross section results with respect to jet mass for groomed jets, p_T 460-550 GeV.

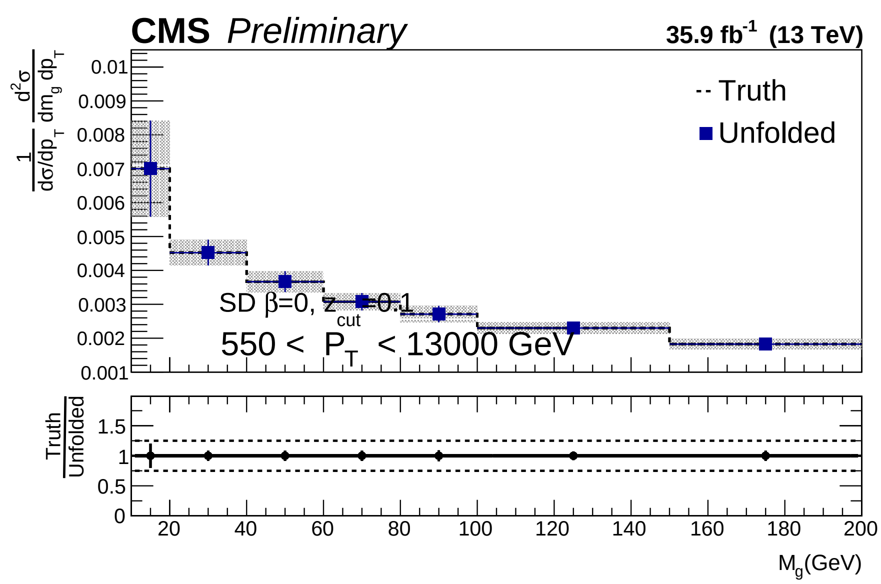

Figure 114: Normalized cross section results with respect to jet mass for groomed jets, p_T 550-13000 GeV.

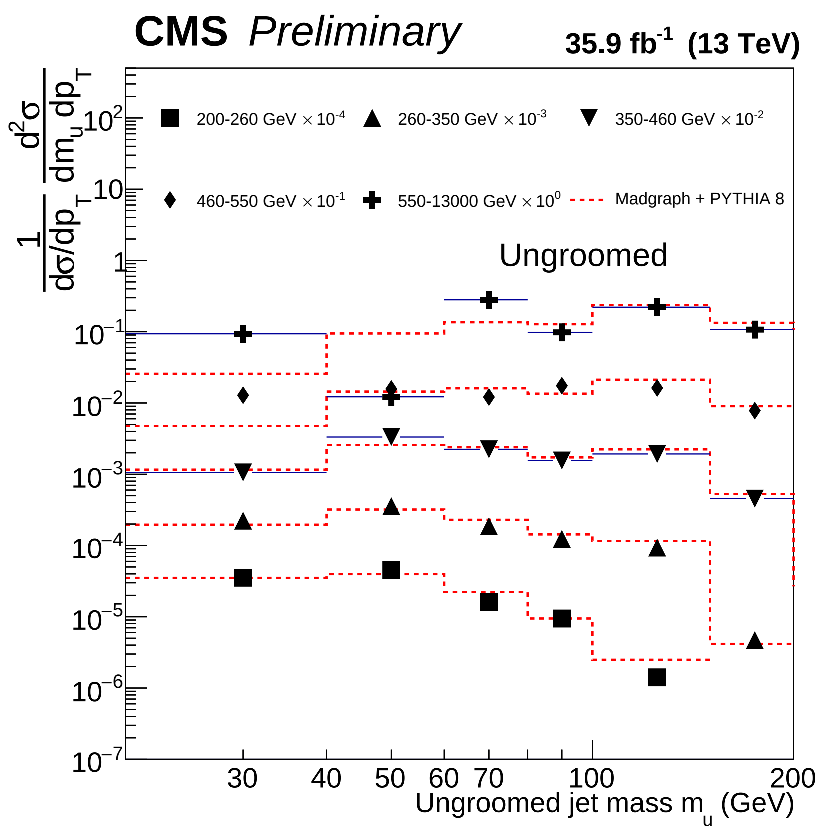

Figure 115: Results for ungroomed reconstructed unfolding with jet mass.

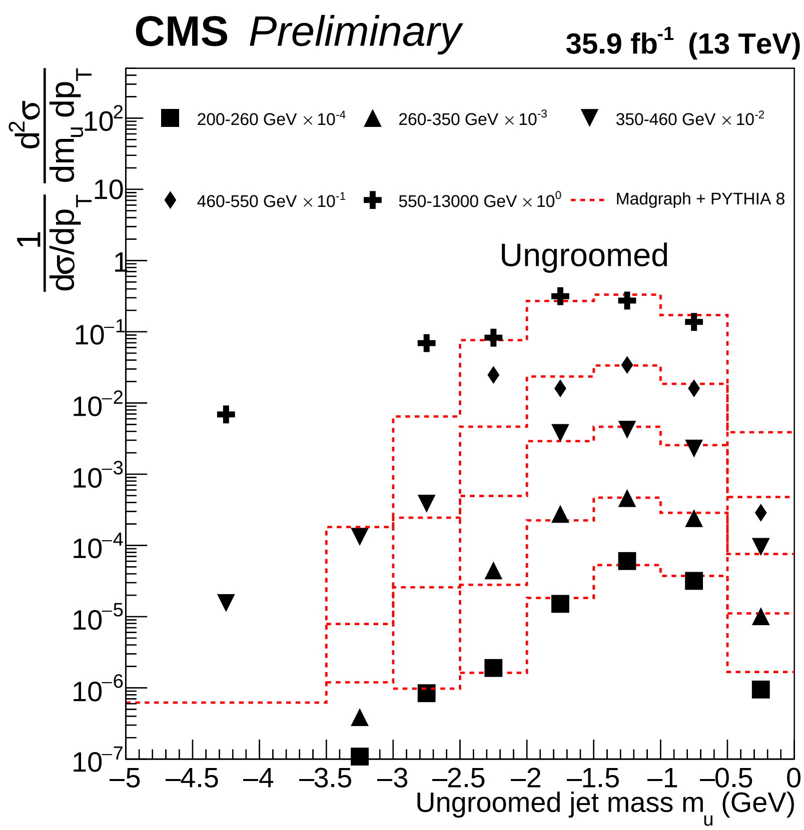

Figure 116: Results for ungroomed reconstructed unfolding with jet dimensionless mass.

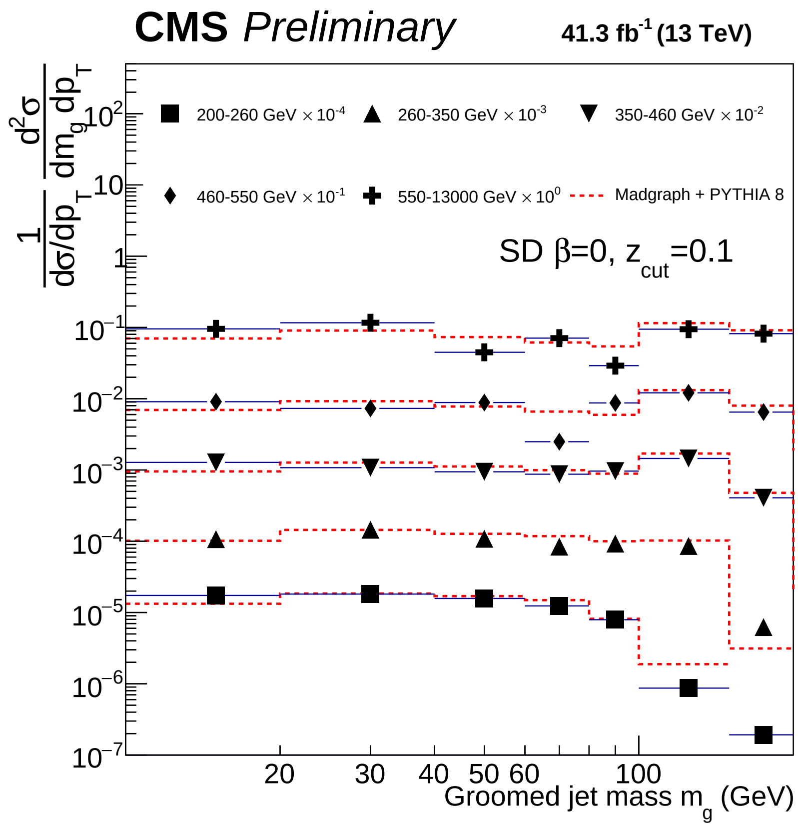

Figure 117: Results for groomed reconstructed unfolding with jet mass.

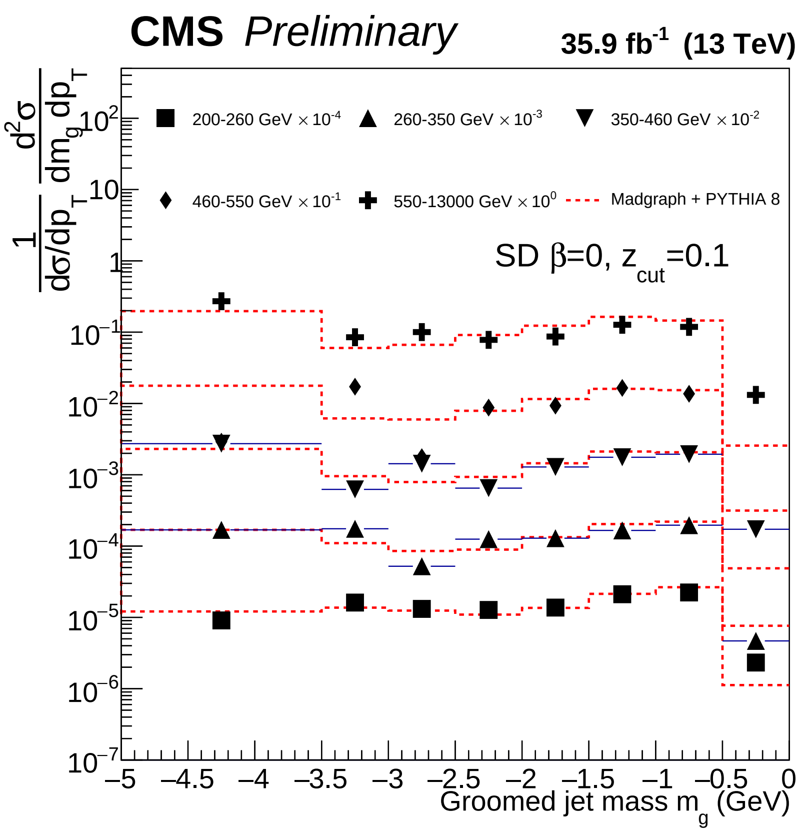

Figure 118: Results for groomed reconstructed unfolding with jet dimensionless mass.

The resulting correlation matrix without systematic uncertainties for the ungroomed jets is shown in Fig. 119, and for the groomed jets is shown in Fig. 121. The same figures with systematic uncertainties are shown for the ungroomed jets in Fig. 123, and for the groomed jets is shown in Fig. 125.

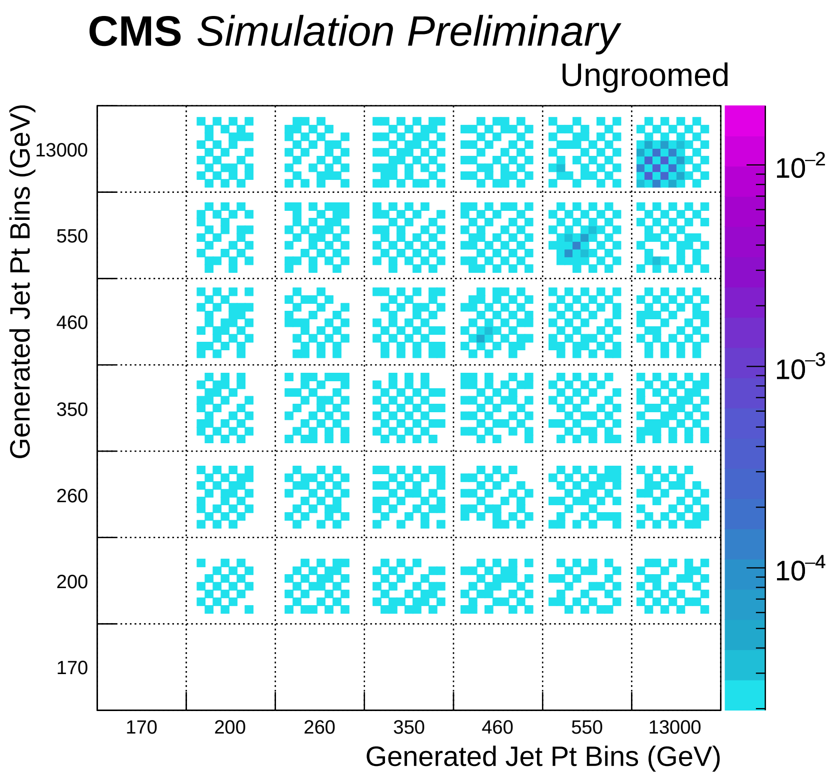

Figure 119: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,10,20,. . . 200 GeV), while the p_T is indexed by the major blocks (200,350,. . . 13000 GeV).

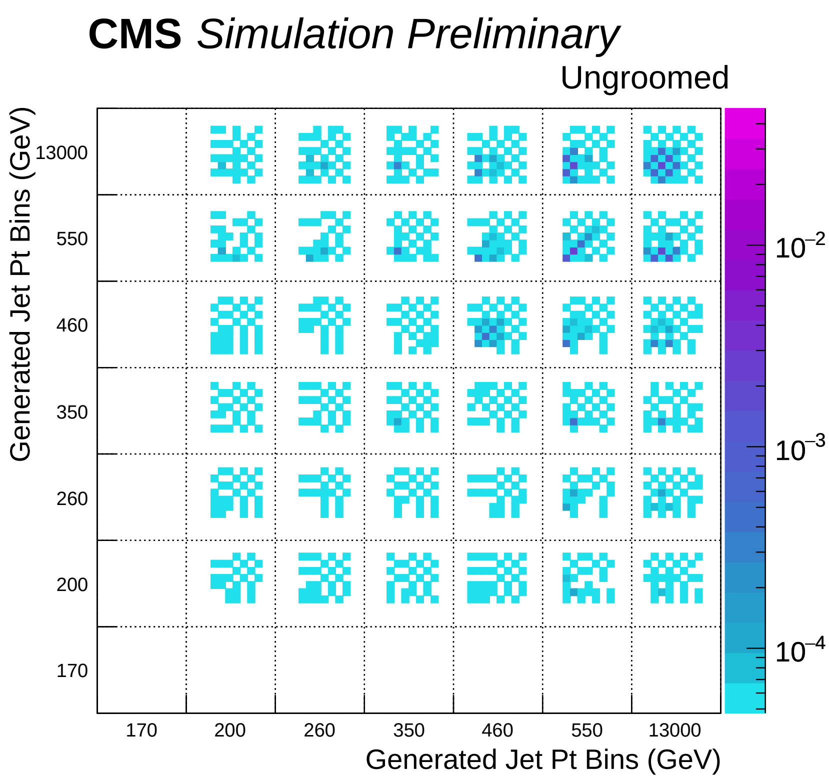

Figure 120: Correlation matrix for ungroomed jets. The bins are set such that the jet dimensionless mass is indexed by the minor blocks (-5, -3.5, -3. . . . 0), while the p_{T} is indexed by the major blocks (200,350,. . . 13000 GeV).

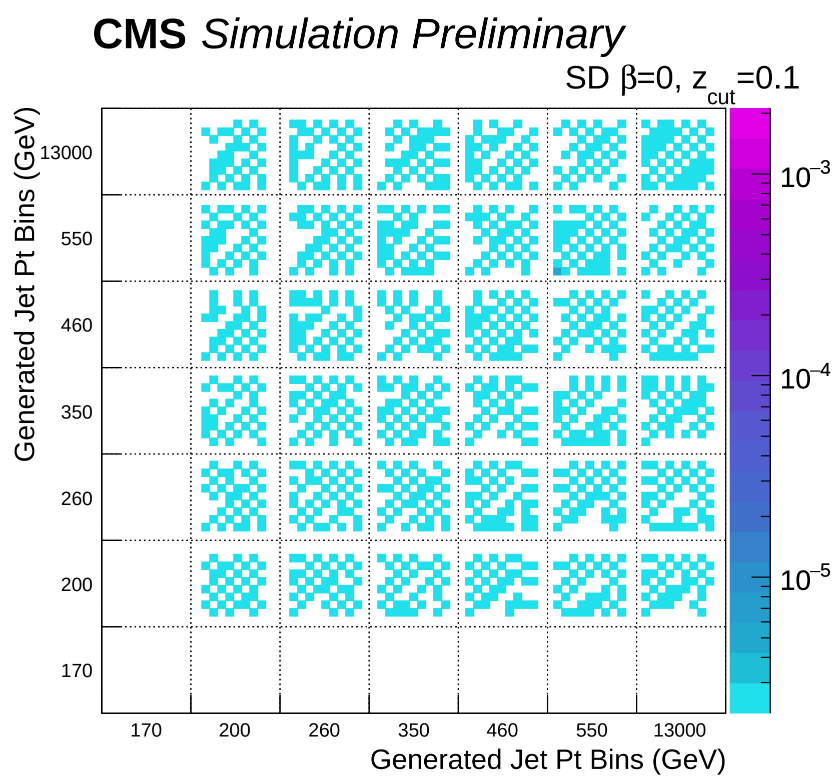

Figure 121: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 760 GeV).

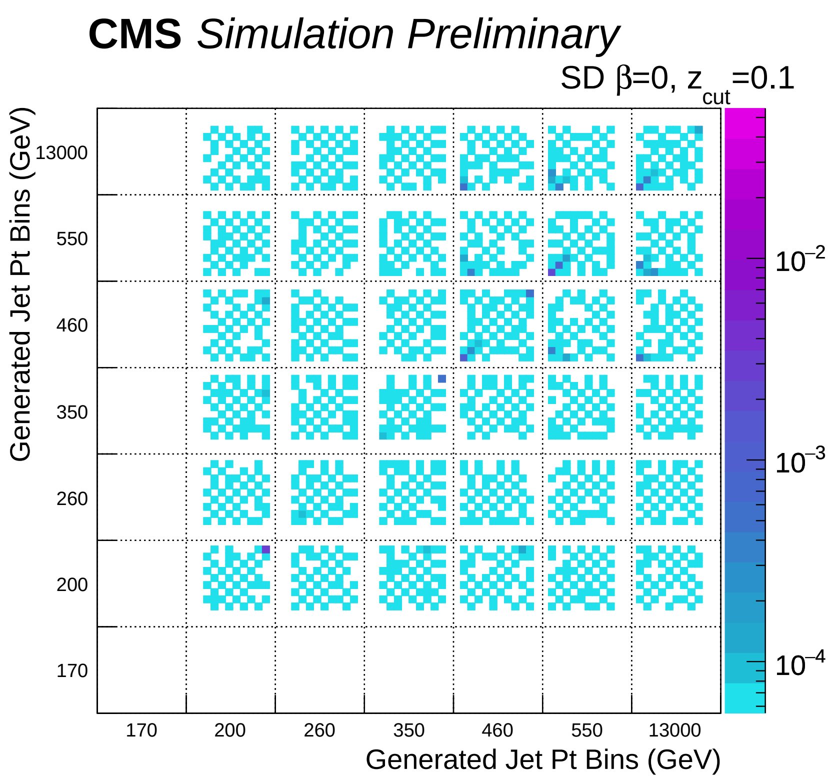

Figure 122: Correlation matrix for groomed jets. The bins are set such that the jet dimensionless mass is indexed by the minor blocks (-5, -3.5, -3. ... 0), while the p_T is indexed by the major blocks (200,340,... 760 GeV).

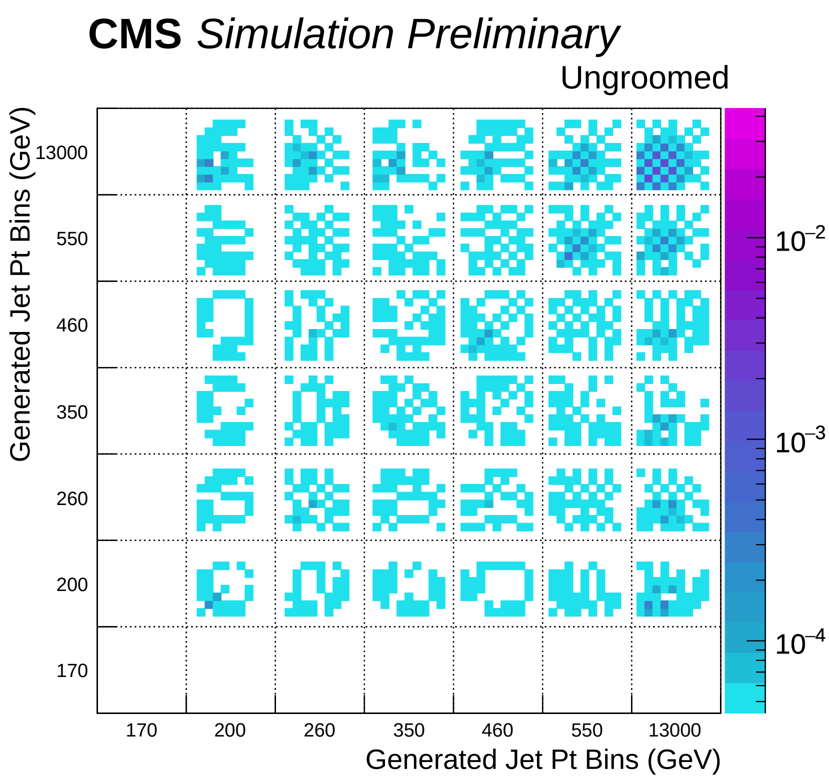

Figure 123: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_{T} is indexed by the major blocks (200,260,. . . 13000 GeV).

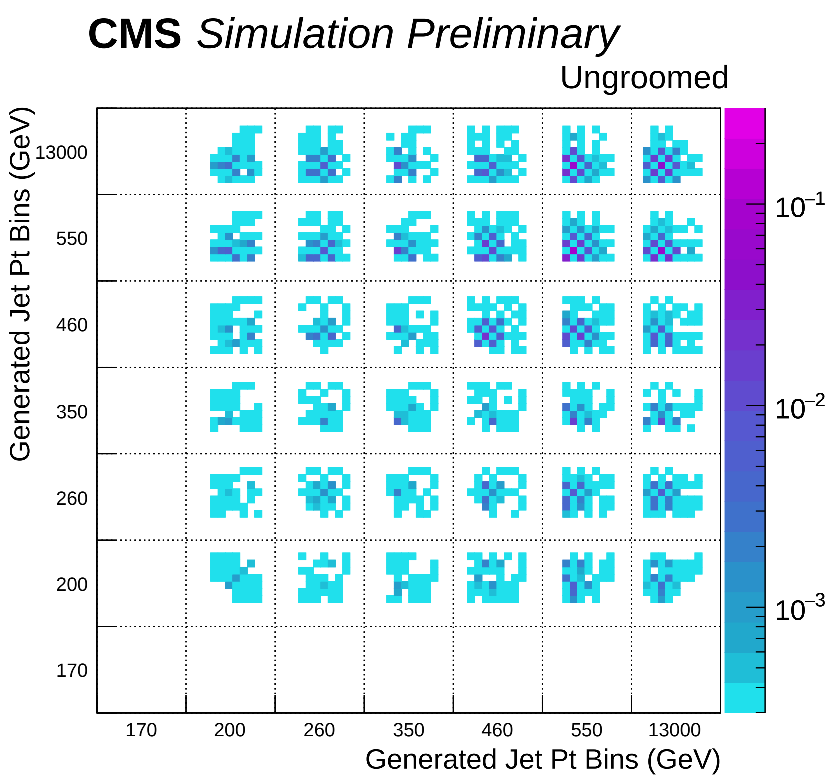

Figure 124: Correlation matrix for ungroomed jets. The bins are set such that the jet dimensionless mass is indexed by the minor blocks (-5, -3.5, -3. . . . 0), while the p_T is indexed by the major blocks (200,260,. . . 13000 GeV).

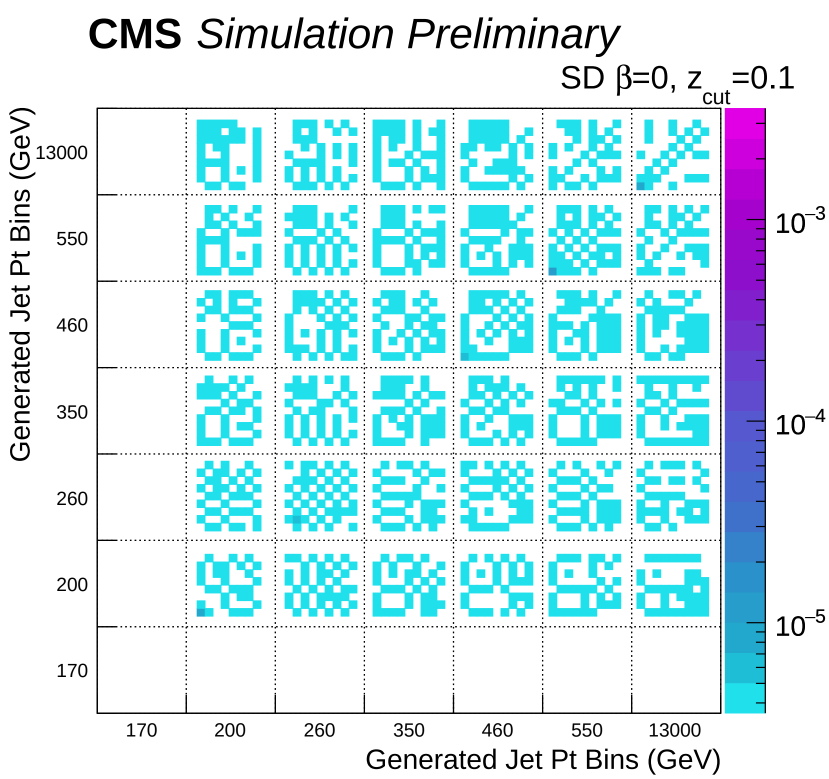

Figure 125: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 0 GeV).

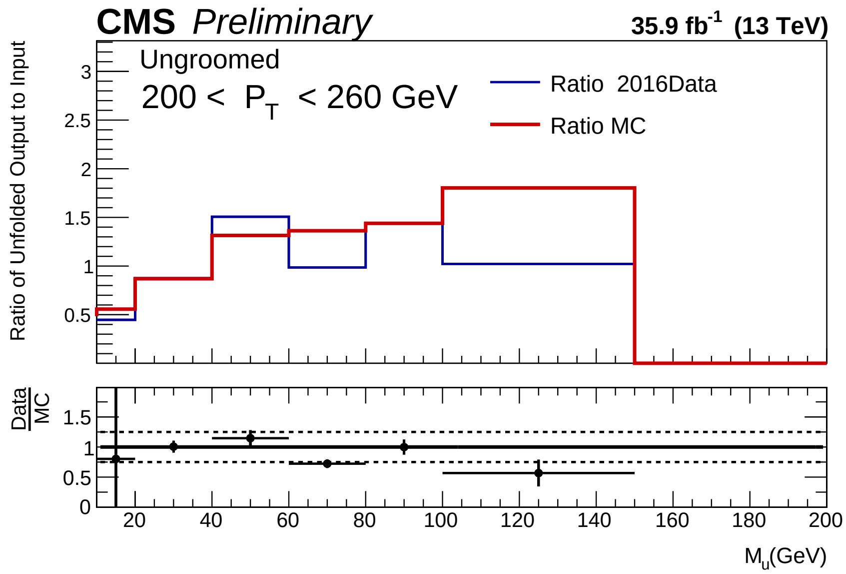

Figure 126: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 200-260 GeV.

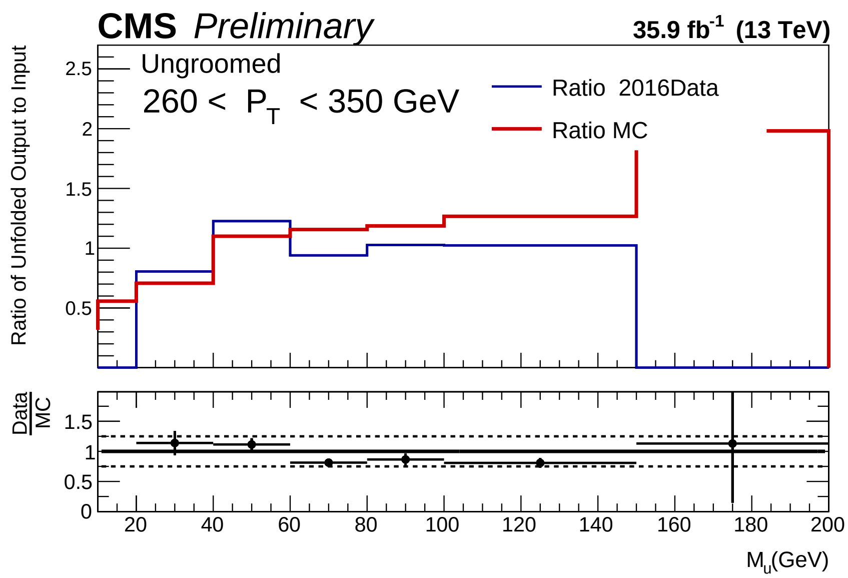

Figure 127: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 260-350 GeV.

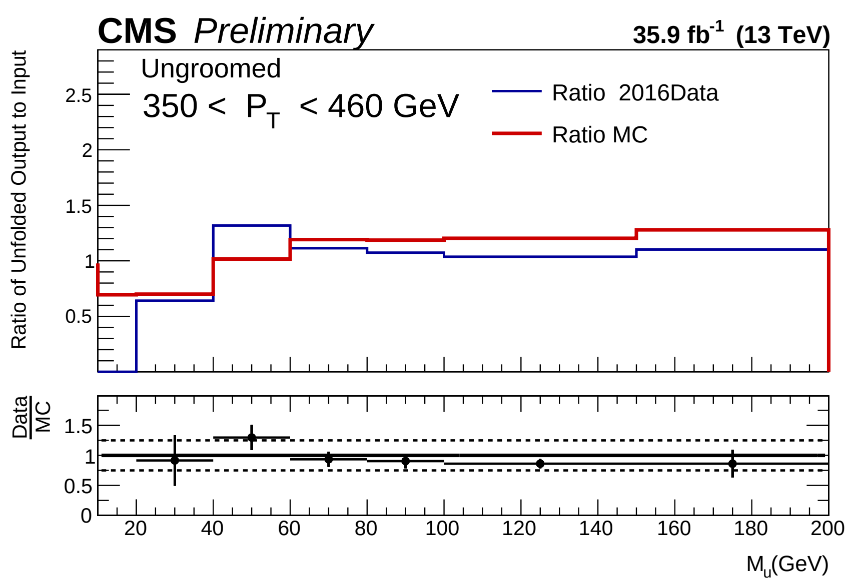

Figure 128: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 350-460 GeV.

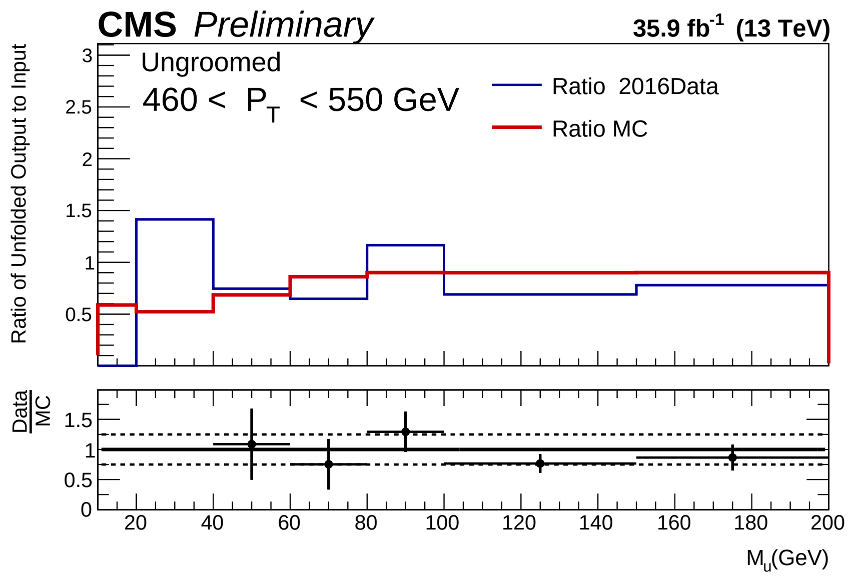

Figure 129: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 460-550 GeV.

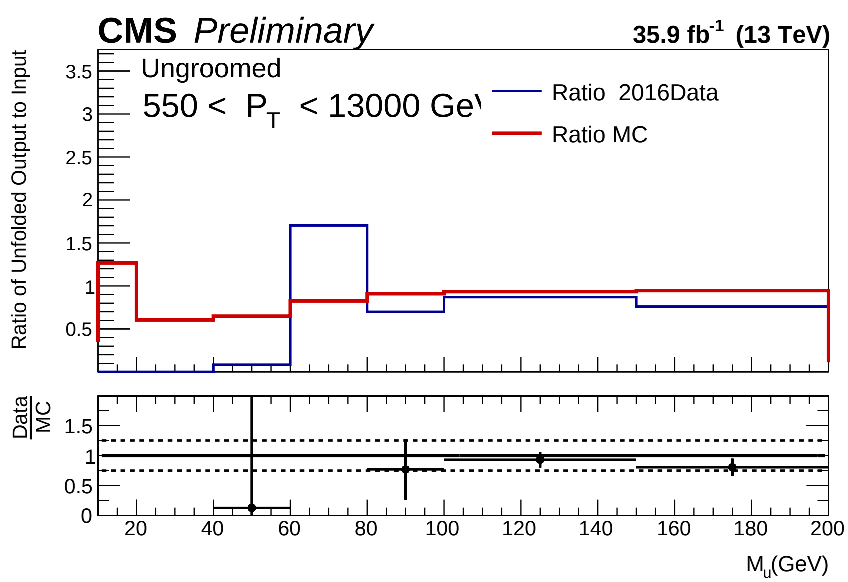

Figure 130: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 550-13000 GeV.

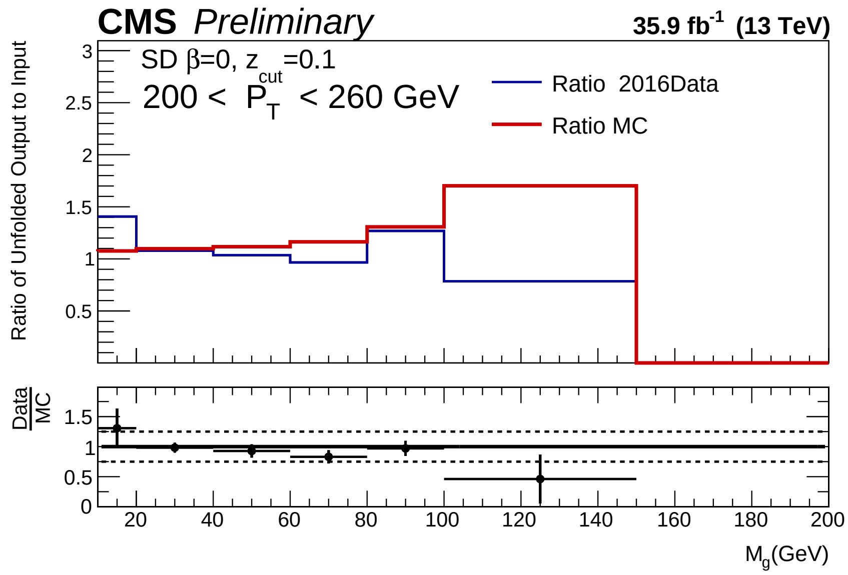

Figure 131: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 200-260 GeV.

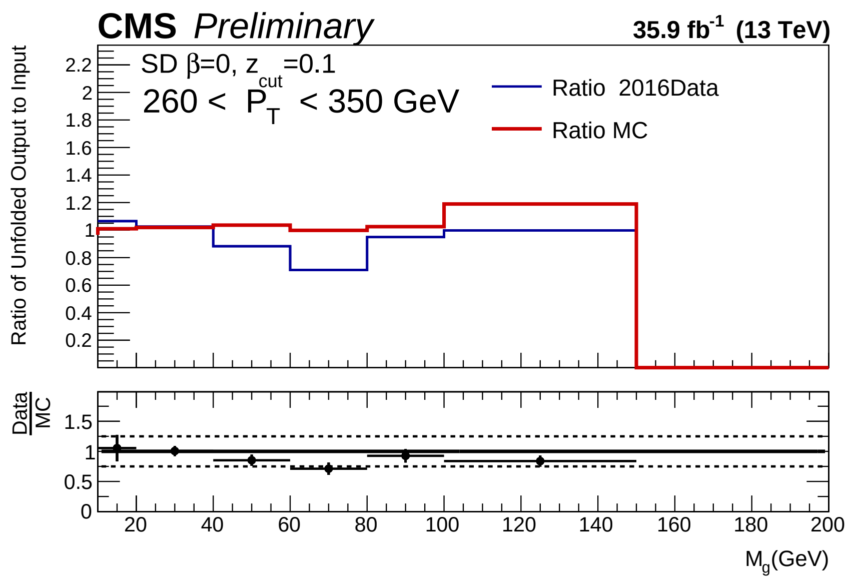

Figure 132: Ratio of unfolded over raw data and MC for groomed jets, p_T 260-350 GeV.

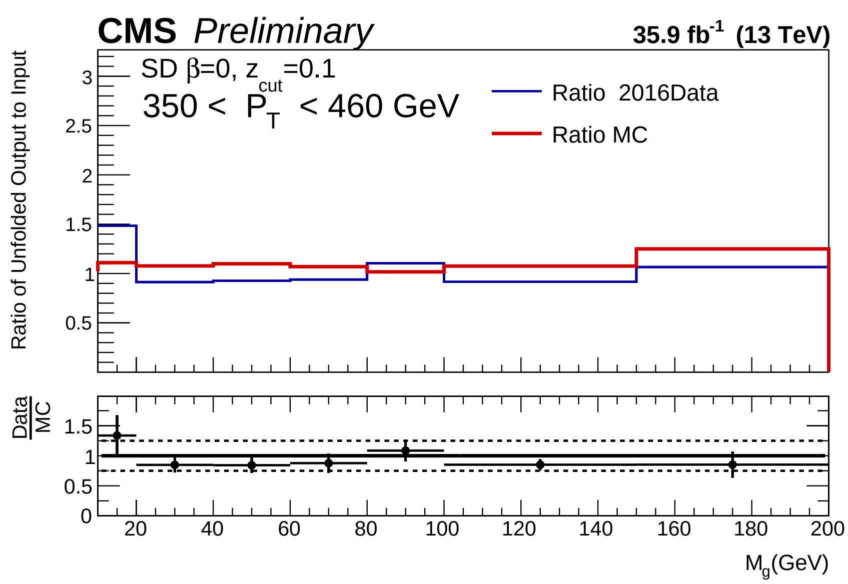

Figure 133: Ratio of unfolded over raw data and MC for groomed jets, p_T 350-460 GeV.

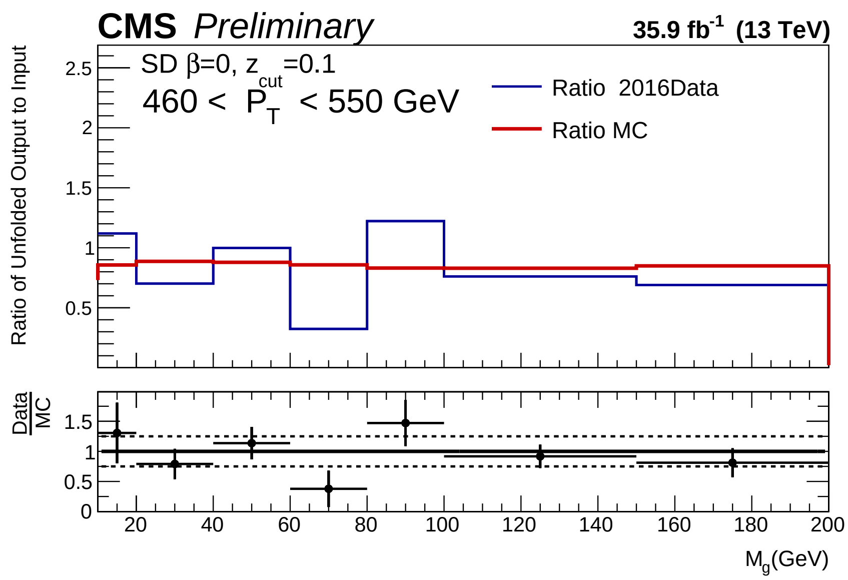

Figure 134: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 460-550 GeV.

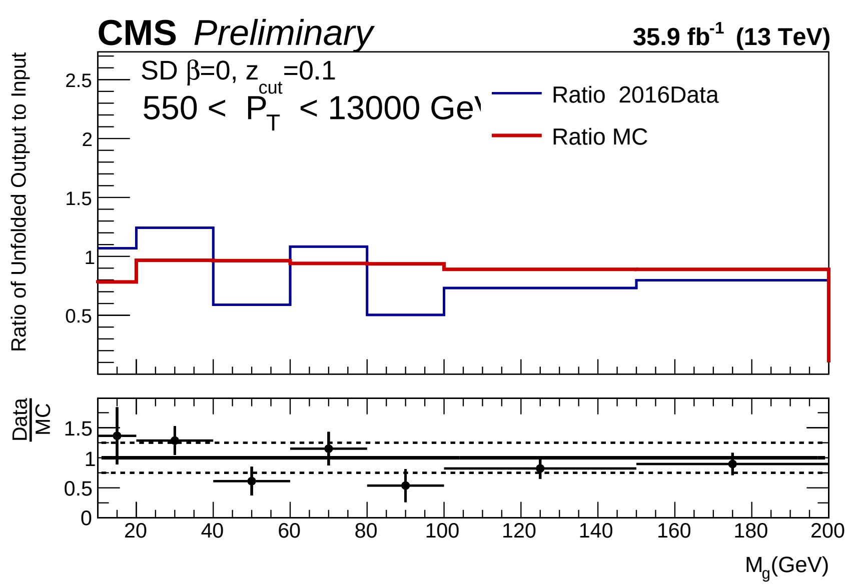

Figure 135: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 550-13000 GeV.

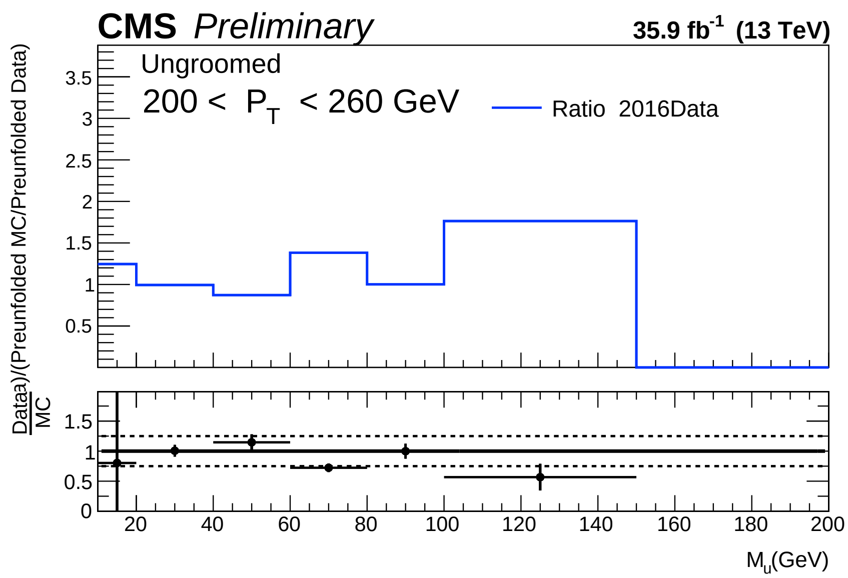

Figure 136: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 200-260 GeV.

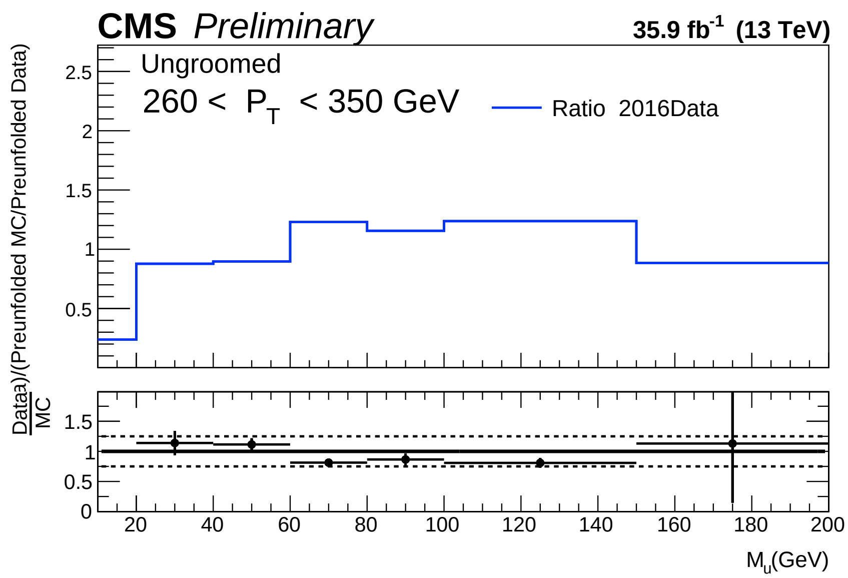

Figure 137: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 260-350 GeV.

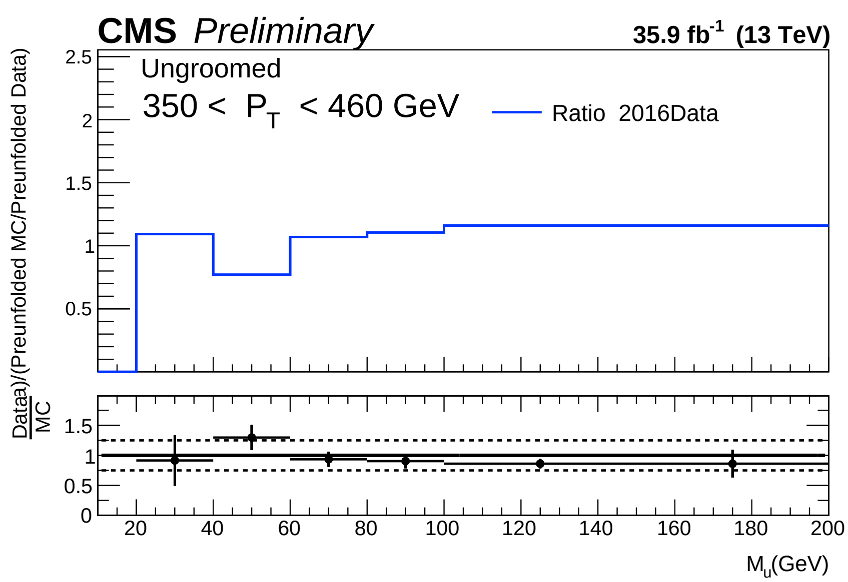

Figure 138: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 350-460 GeV.

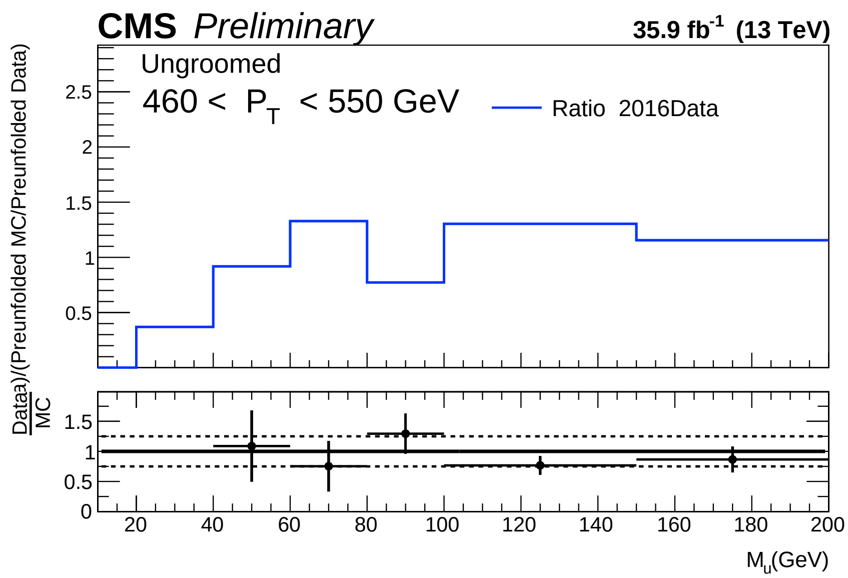

Figure 139: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 460-550 GeV.

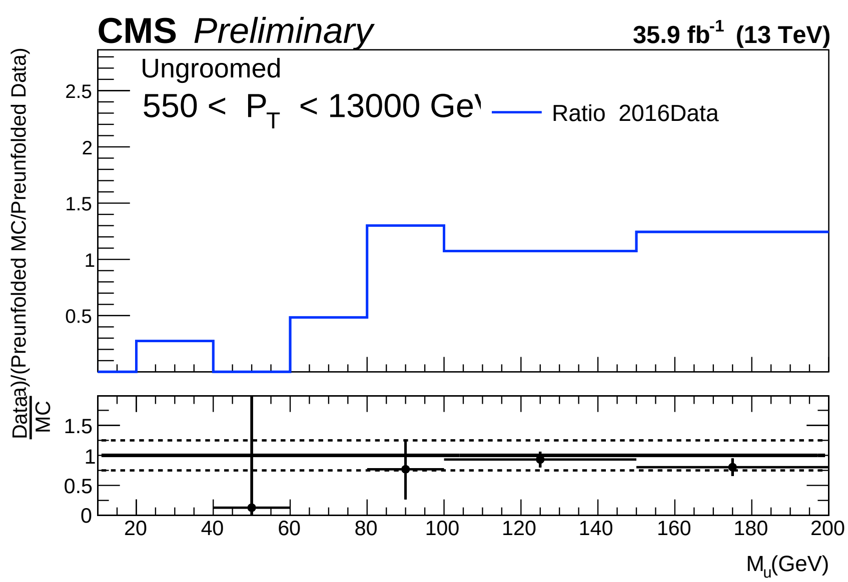

Figure 140: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 550-13000 GeV.

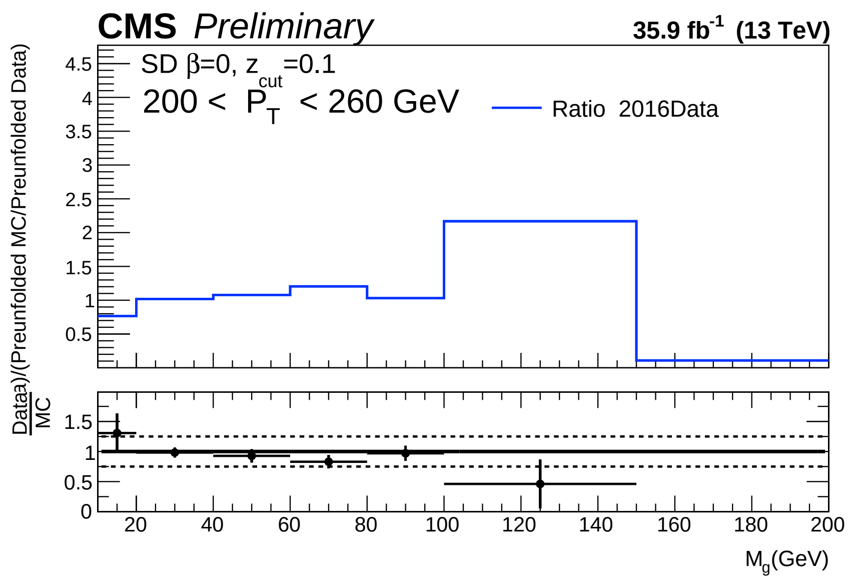

Figure 141: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_T 200-260 GeV.

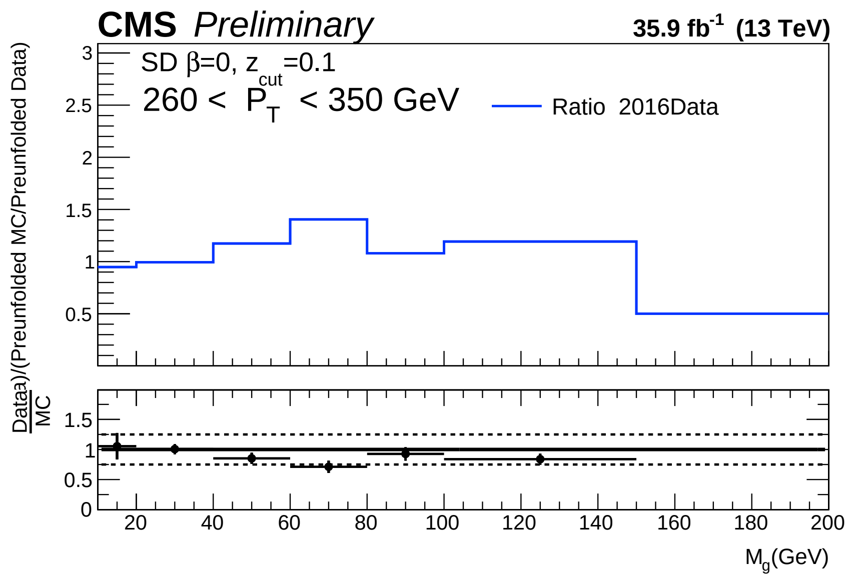

Figure 142: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 260-350 GeV.

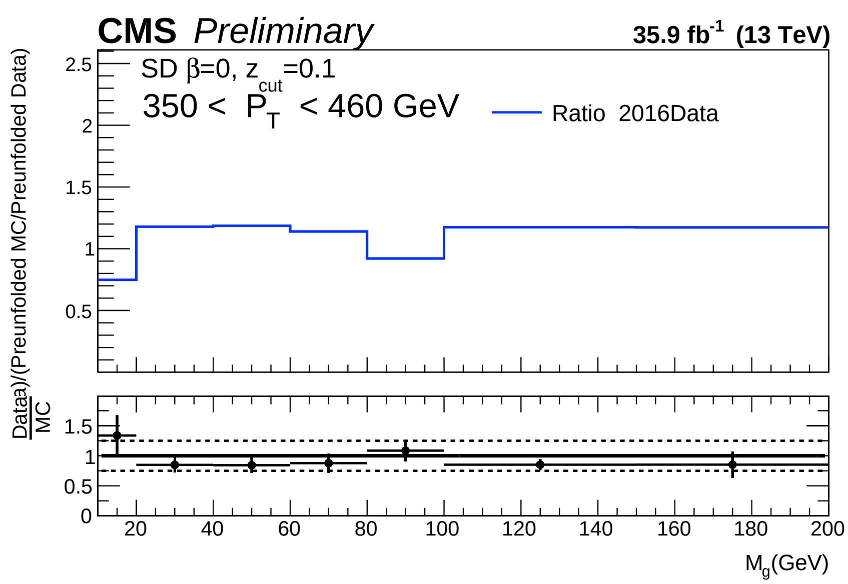

Figure 143: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 350-460 GeV.

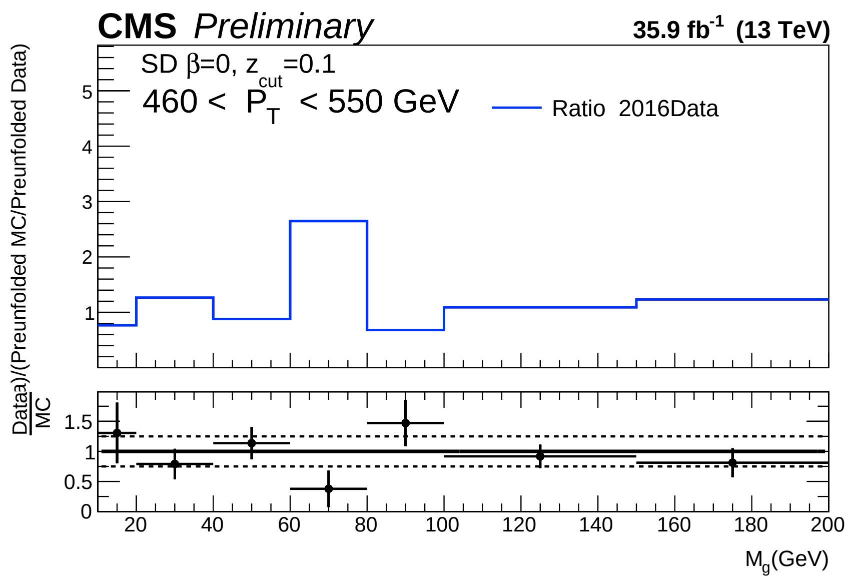

Figure 144: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 460-550 GeV.

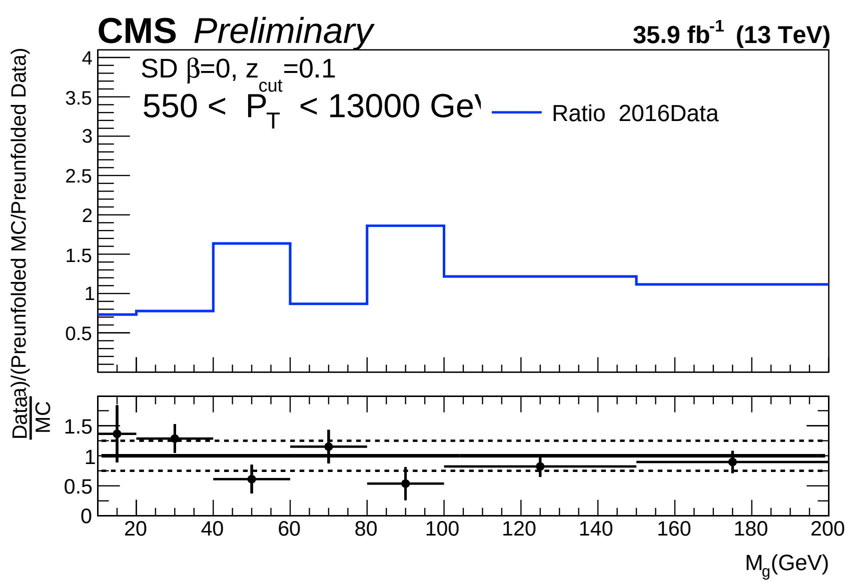

Figure 145: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 550-13000 GeV.

.2 2017 data results

This Appendix shows the distributions from the ”Detector Response” through the ”Results” sections of the main analysis note with only the 2017 data rather than the full Run 2 statistics seen in the main body of the note.

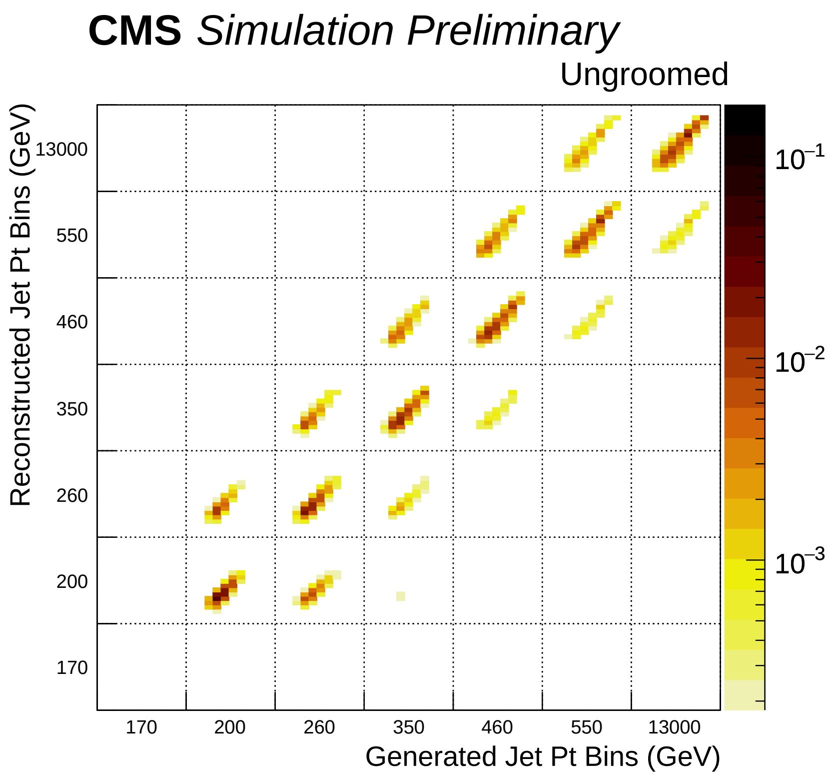

Figure 146: Two-dimensional response matrix for ungroomed jets. The bins are set such that the p_T is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV). While the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

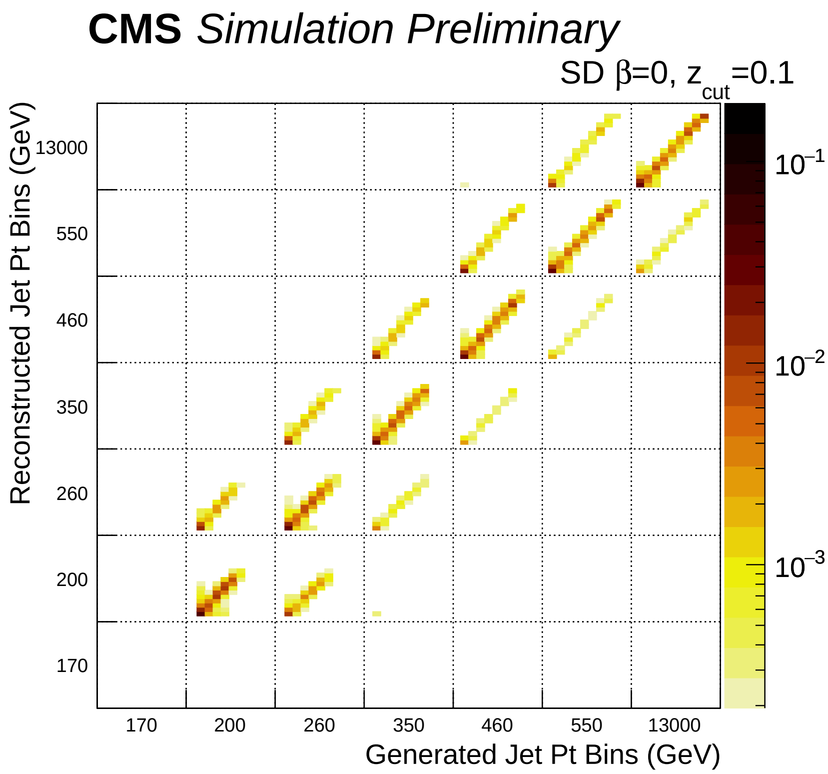

Figure 147: Two-dimensional response matrix for groomed jets $\beta = 0$, $z_{cut} = 0.1$. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV).while the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

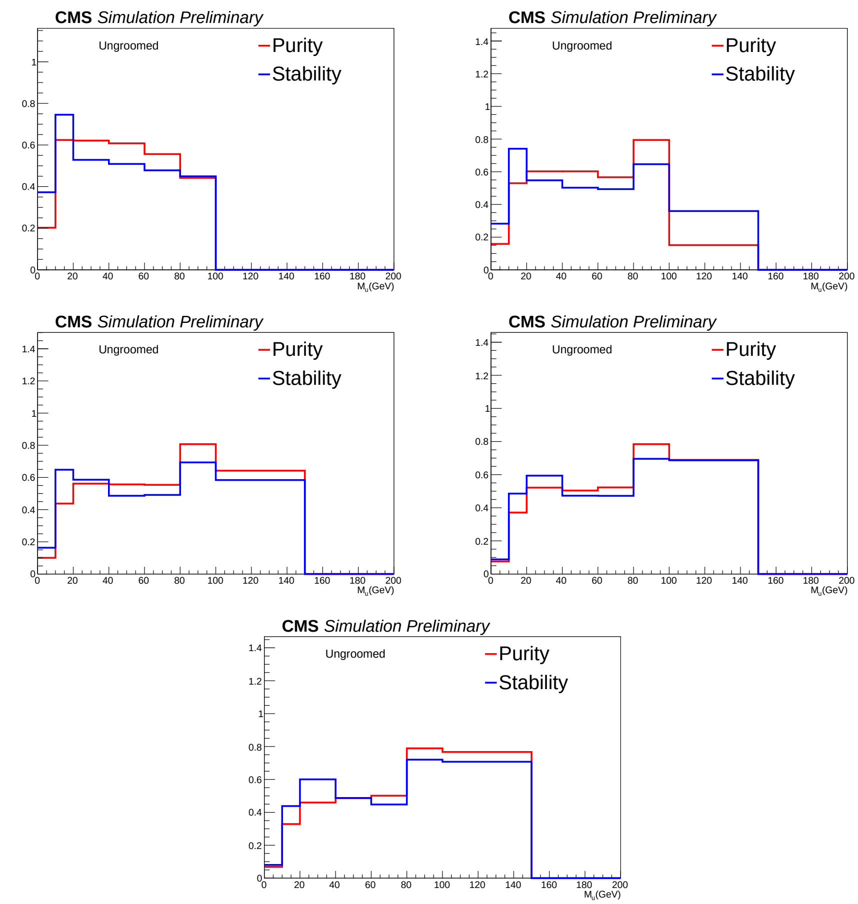

Figure 148: Purity and stability for ungroomed jets in various p_T bins.

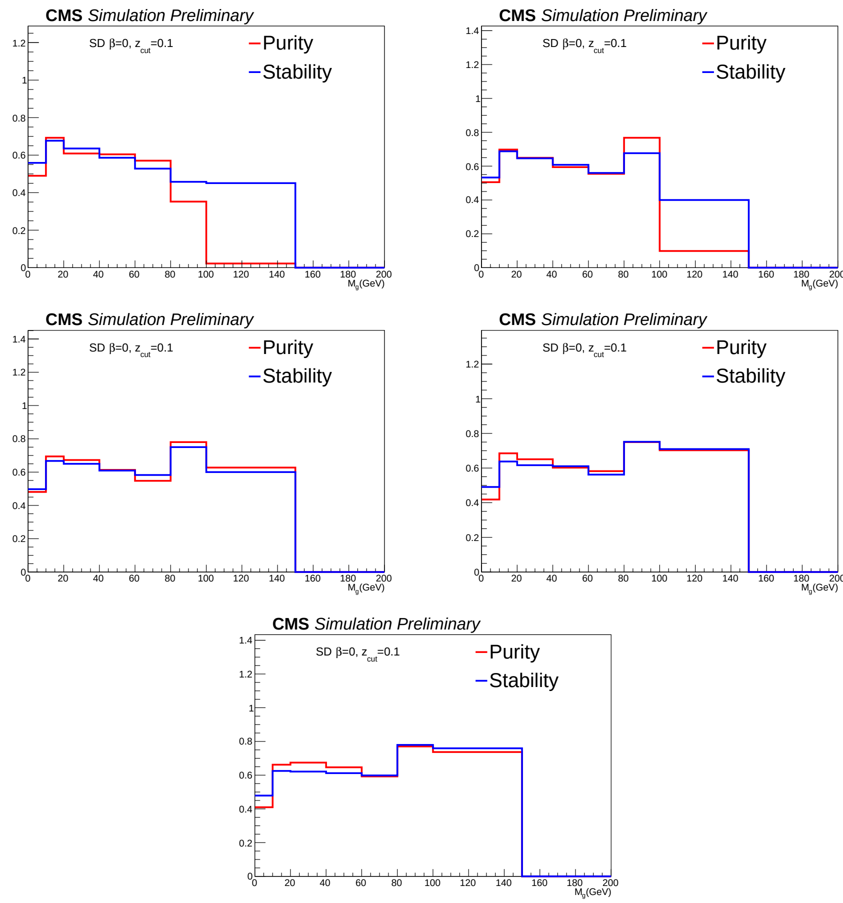

Figure 149: Purity and stability for groomed jets, where the soft-drop criterion was applied with $\beta = 0$ and $z_{cut} = 0.1$, in various p_{T} bins.

Figure-150 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_{T} bin for the normalized cross section. Figure-151 shows the same distributions for the raw dimensionless mass.

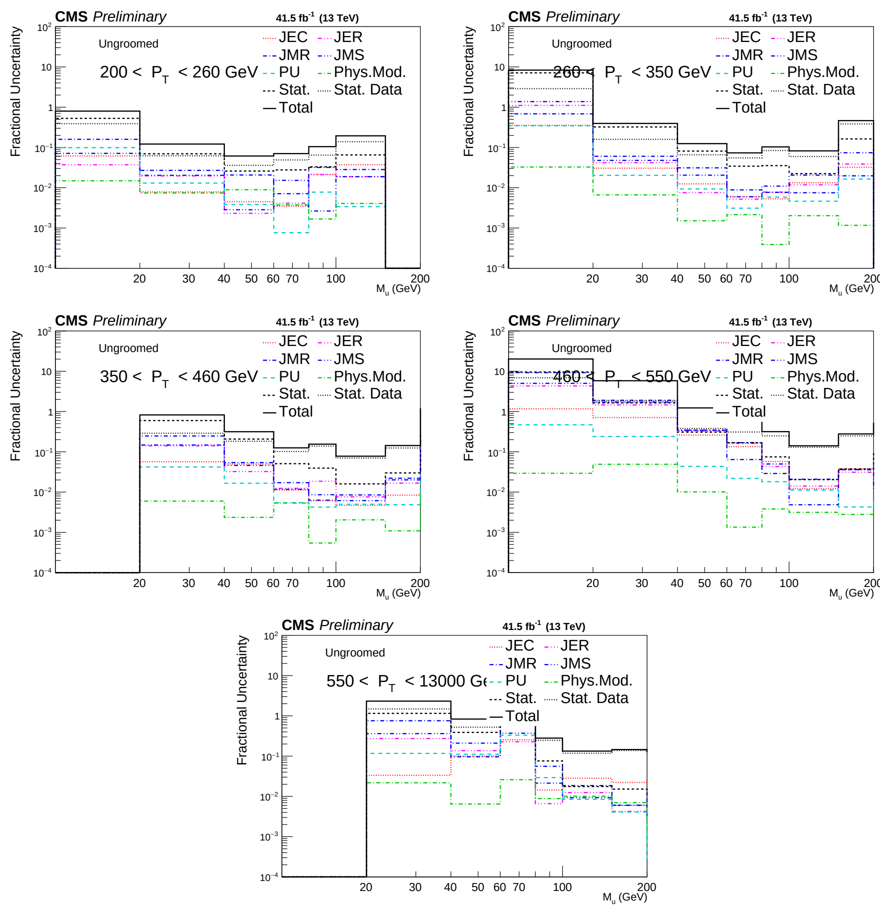

Figure 150: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_{T} bins for the raw jet mass.

Figure-152 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_{T} bin for the normalized cross section. Figure-153 shows the same distributions for the groomed dimensionless mass.

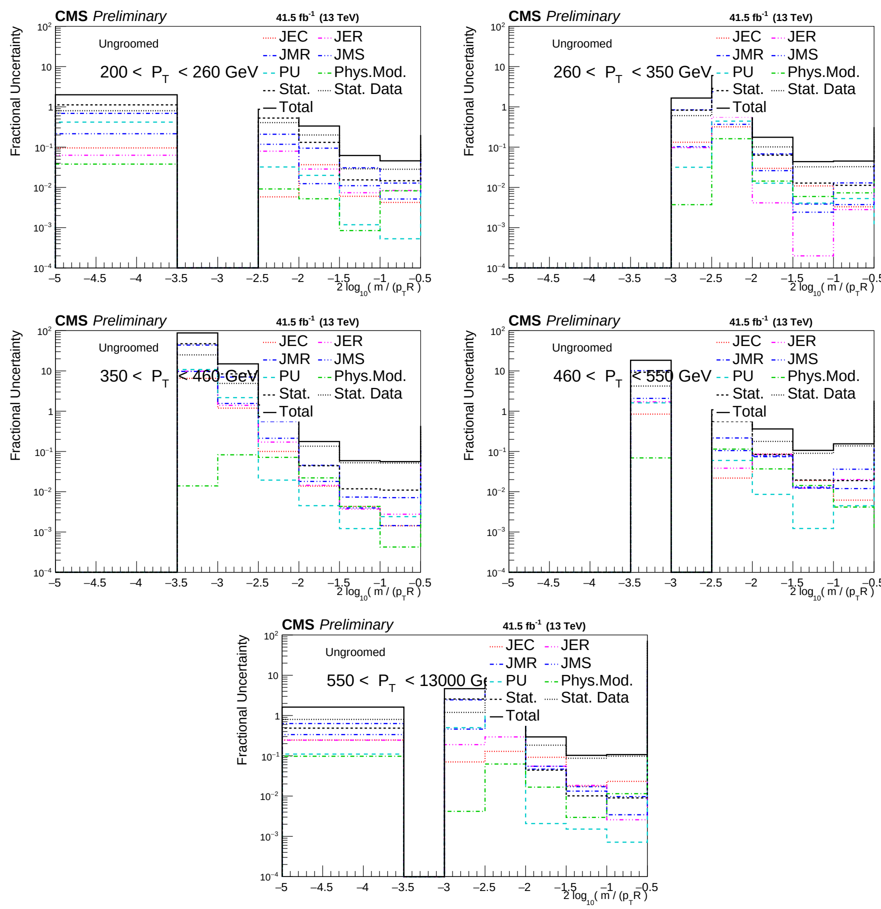

Figure 151: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the raw dimensionless jet mass.

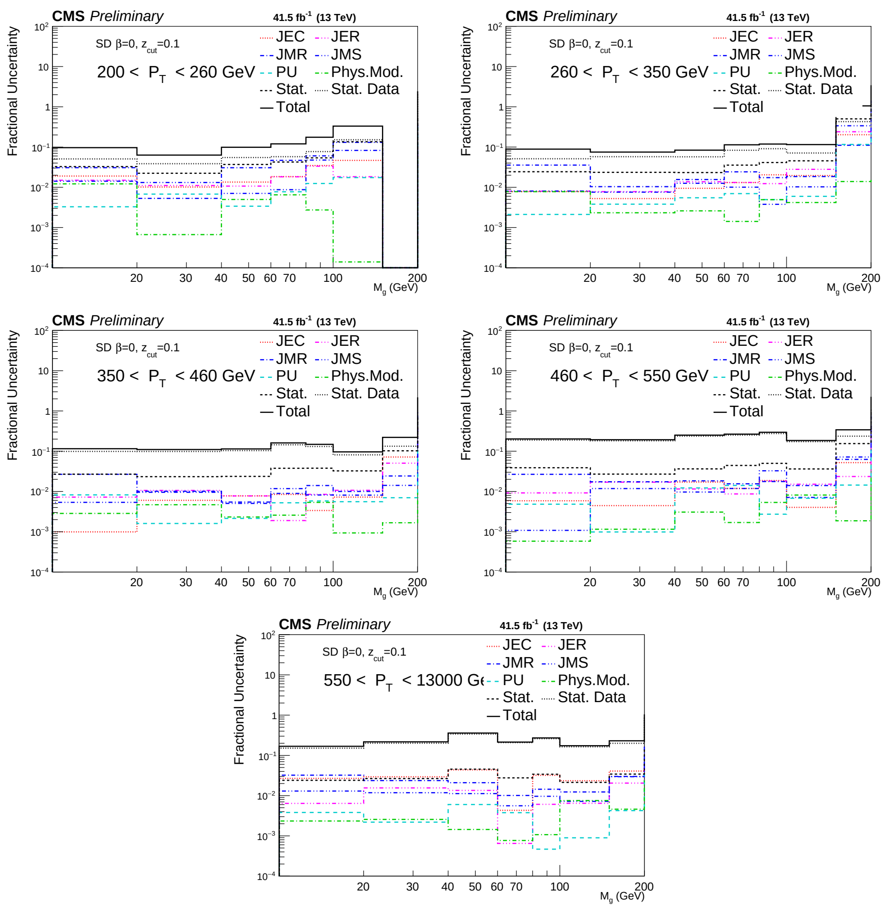

Figure 152: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_{T} bins for the jet mass after grooming with default Soft-Drop.

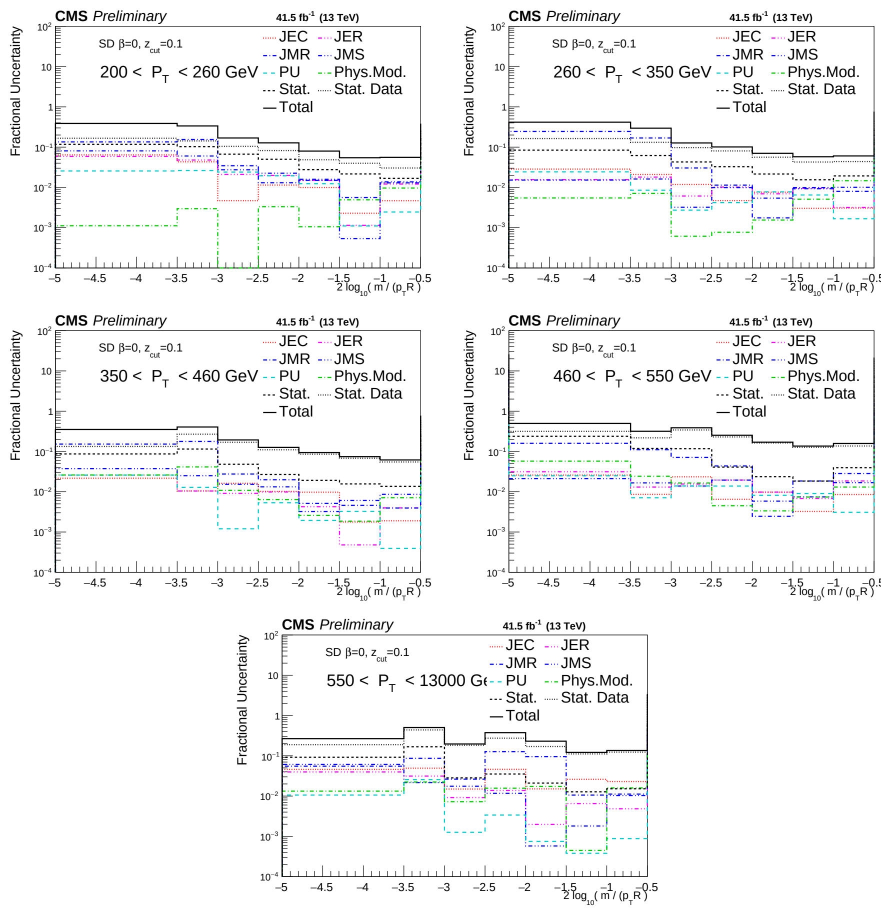

Figure 153: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the jet mass after grooming with default Soft-Drop.

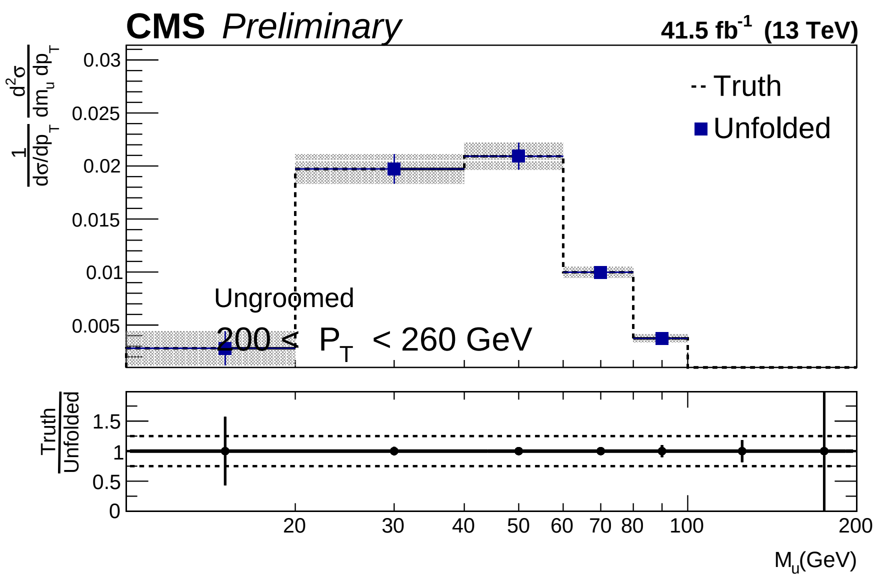

Figure 154: Closure test of ungroomed reconstructed Monte Carlo, p_T 200-260 GeV.

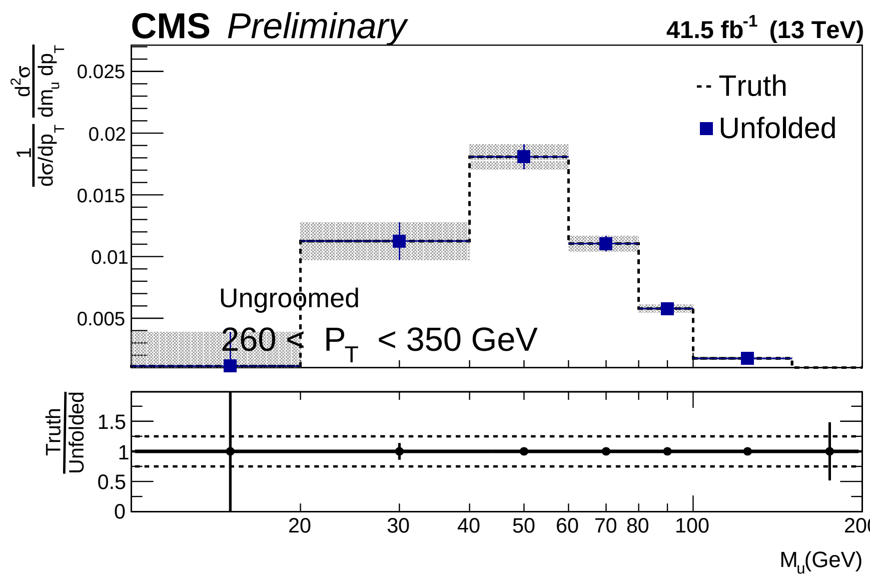

Figure 155: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

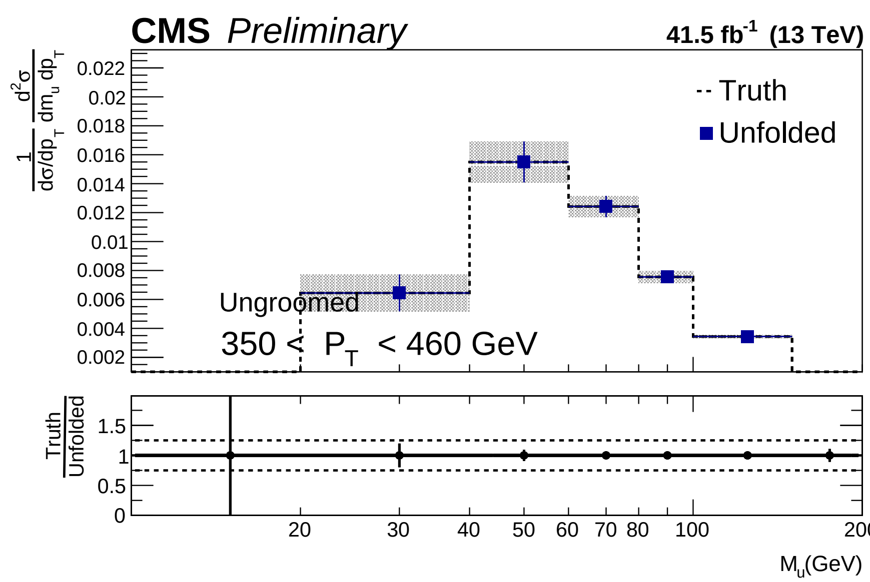

Figure 156: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

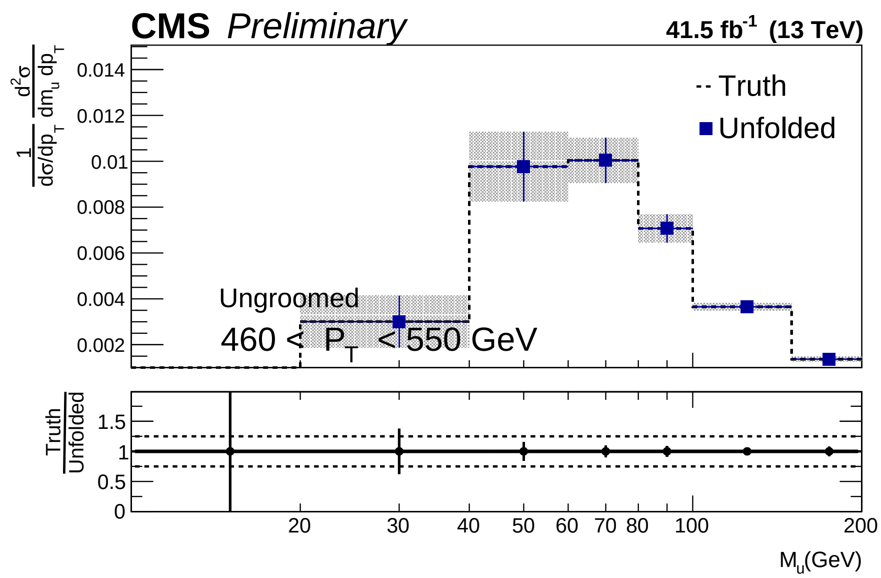

Figure 157: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

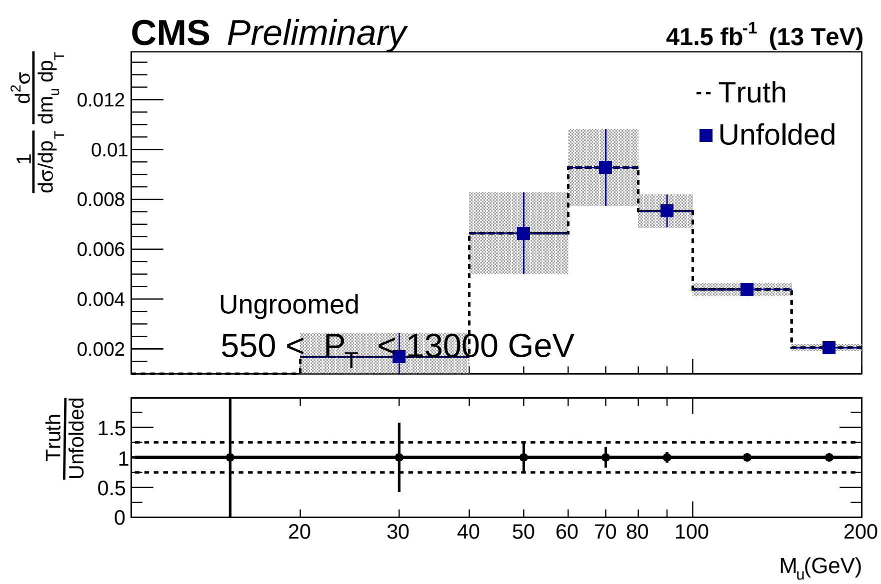

Figure 158: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

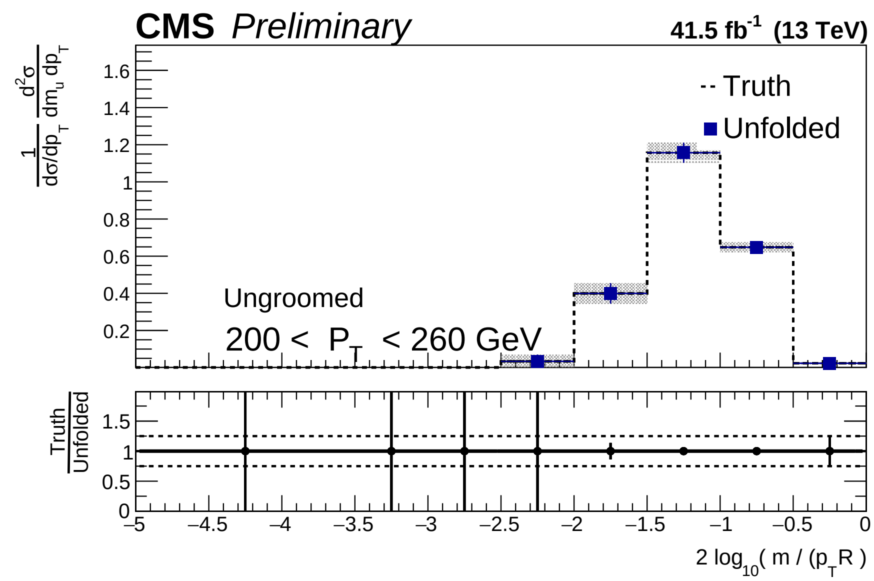

Figure 159: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

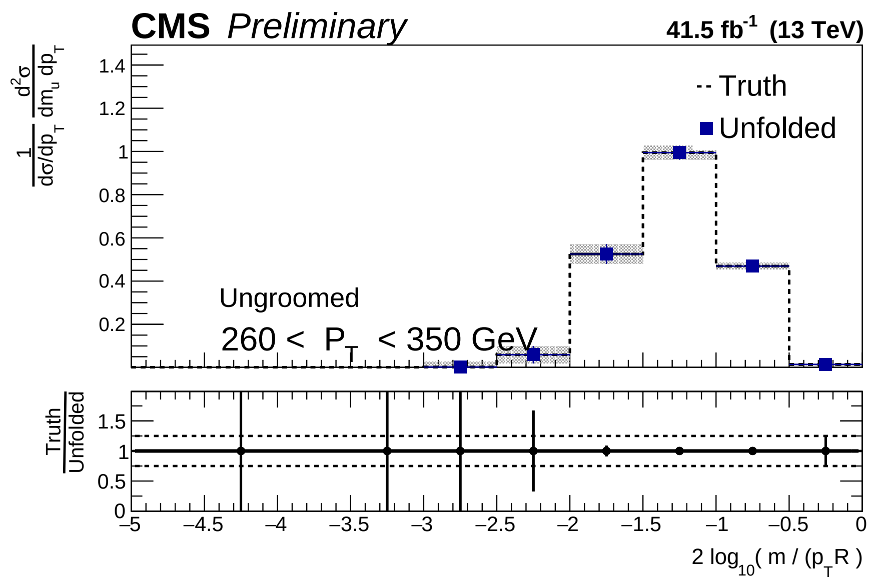

Figure 160: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 260-350 GeV.

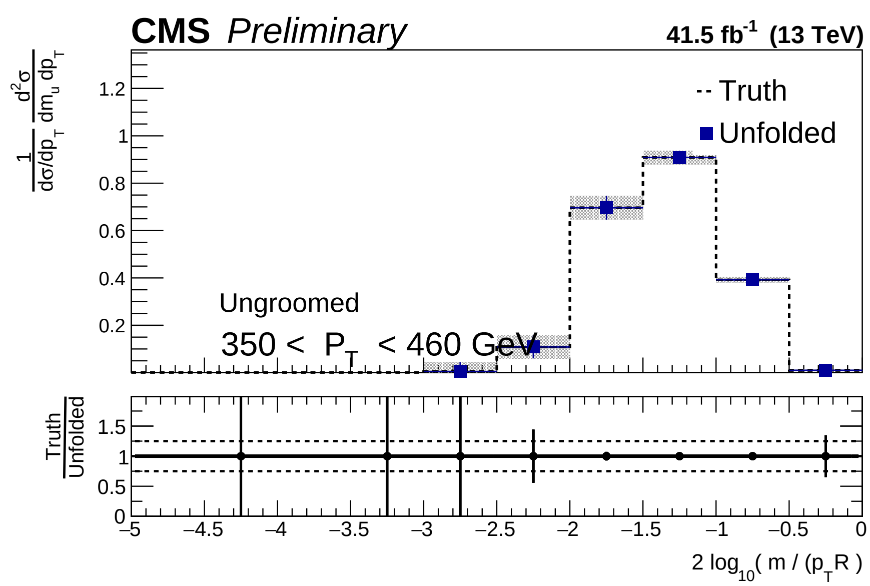

Figure 161: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 350-460 GeV.

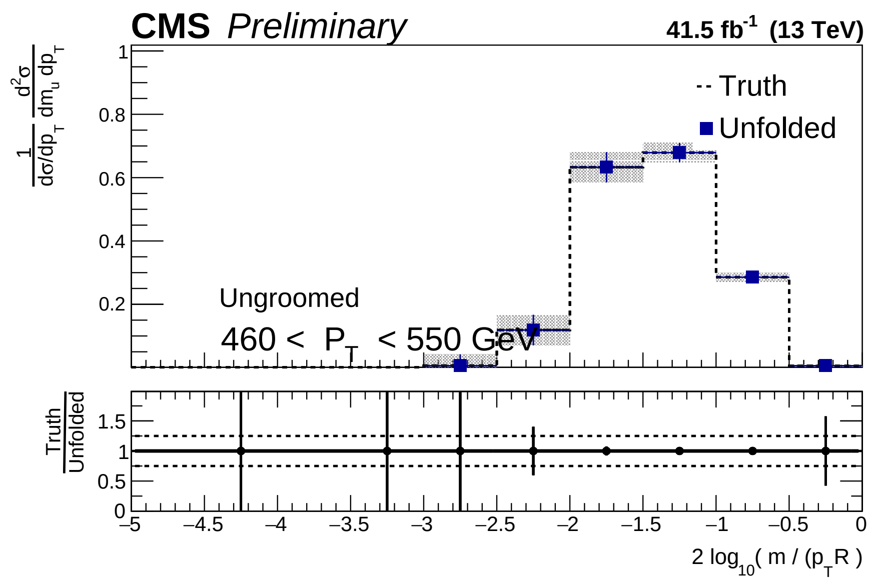

Figure 162: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 460-550 GeV.

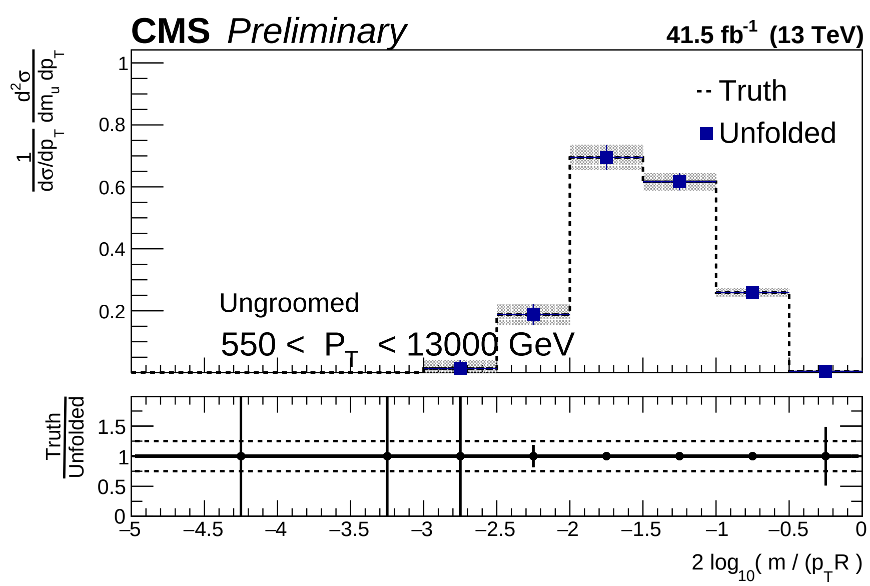

Figure 163: Dimensionless mass closure test of ungroomed reconstructed Monte Carlo, p_T 550-13000 GeV.

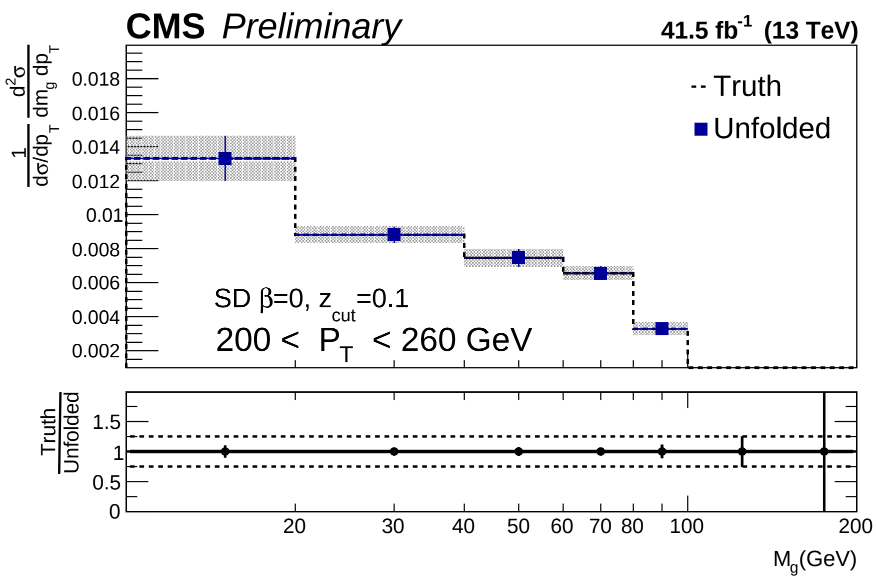

Figure 164: Closure test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

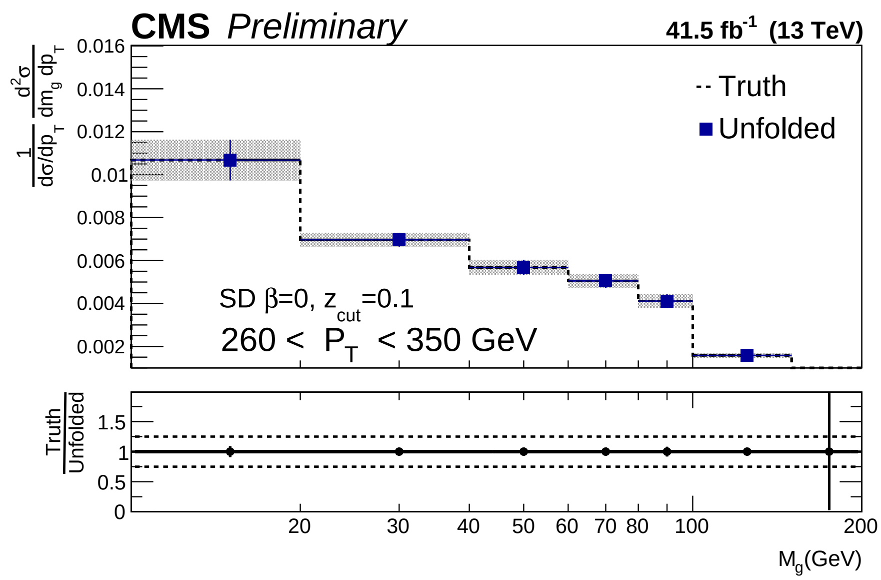

Figure 165: Closure test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

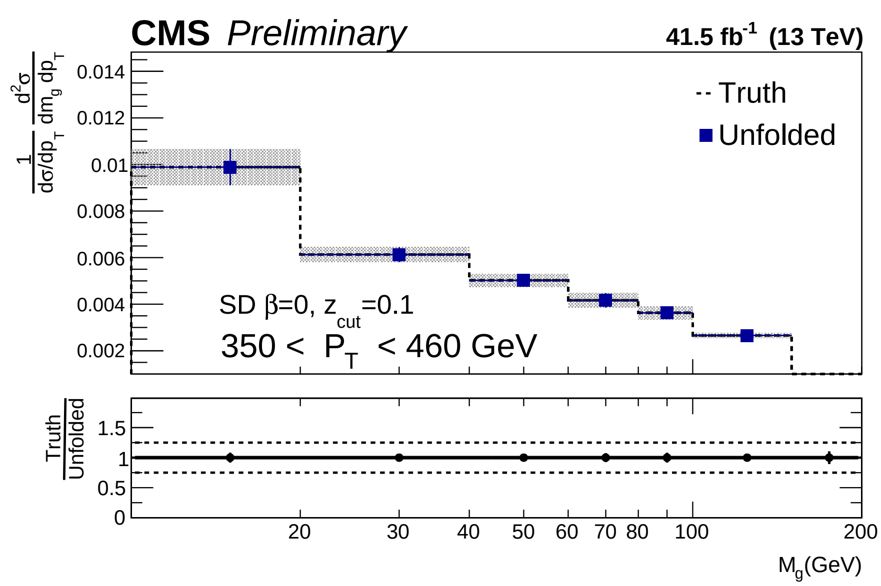

Figure 166: Closure test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

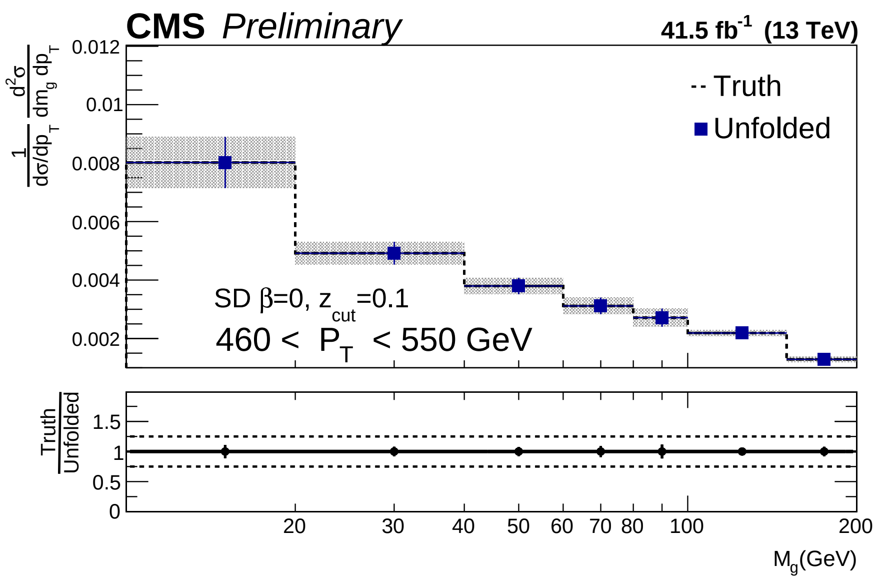

Figure 167: Closure test of groomed reconstructed Monte Carlo, p_T 460-550 GeV.

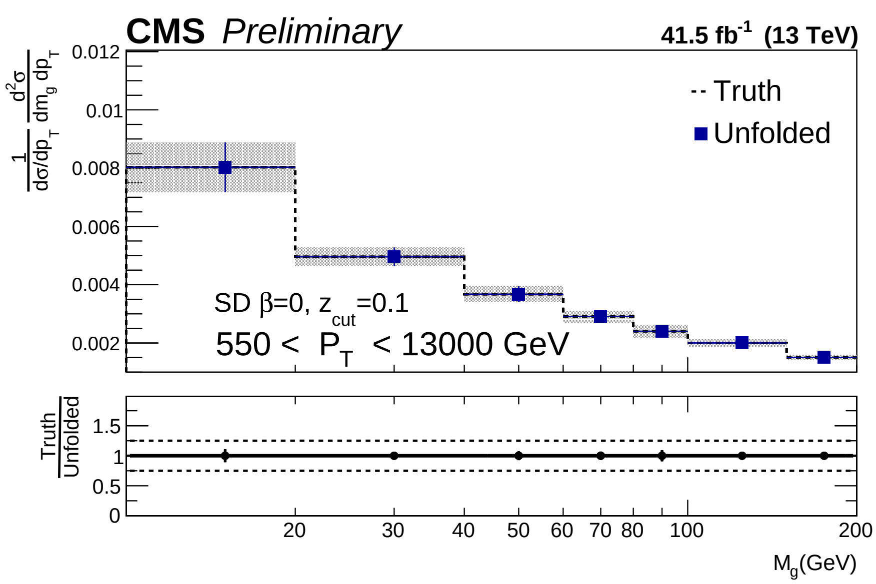

Figure 168: Closure test of groomed reconstructed Monte Carlo, p_T 550-13000 GeV.

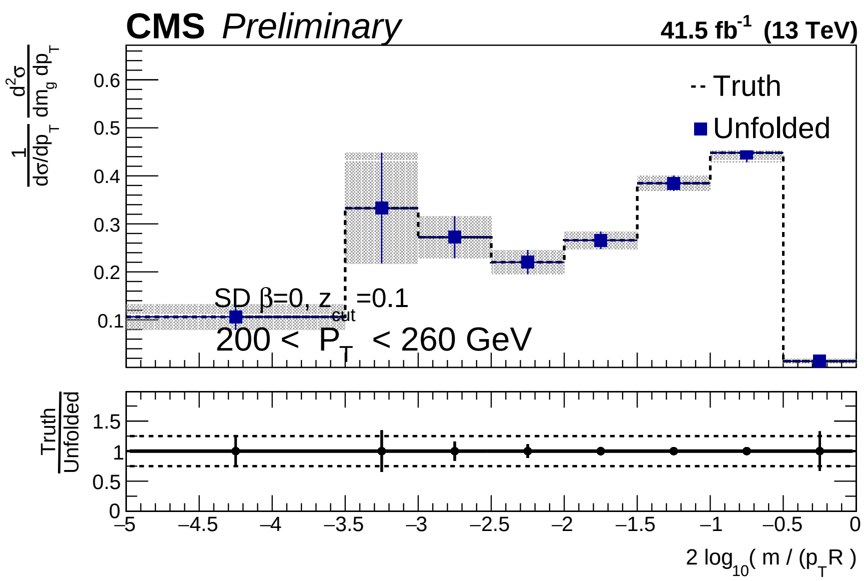

Figure 169: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

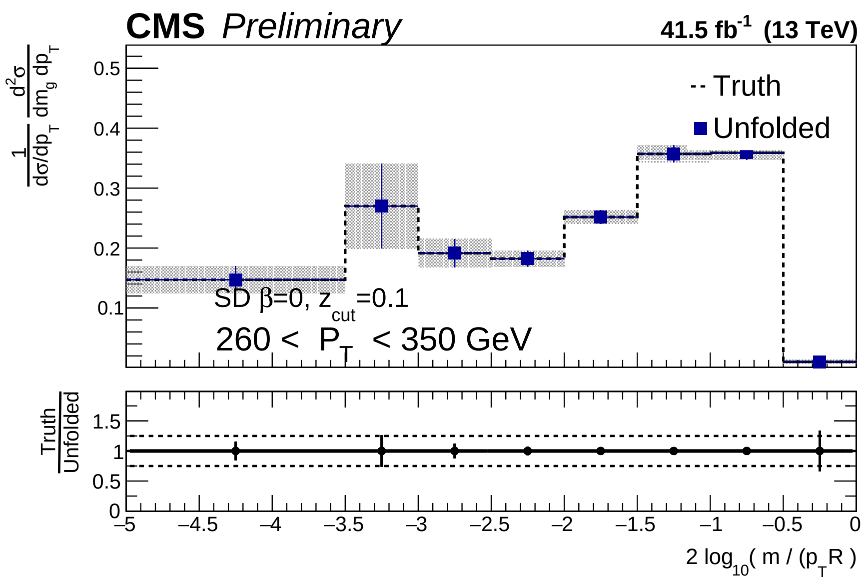

Figure 170: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

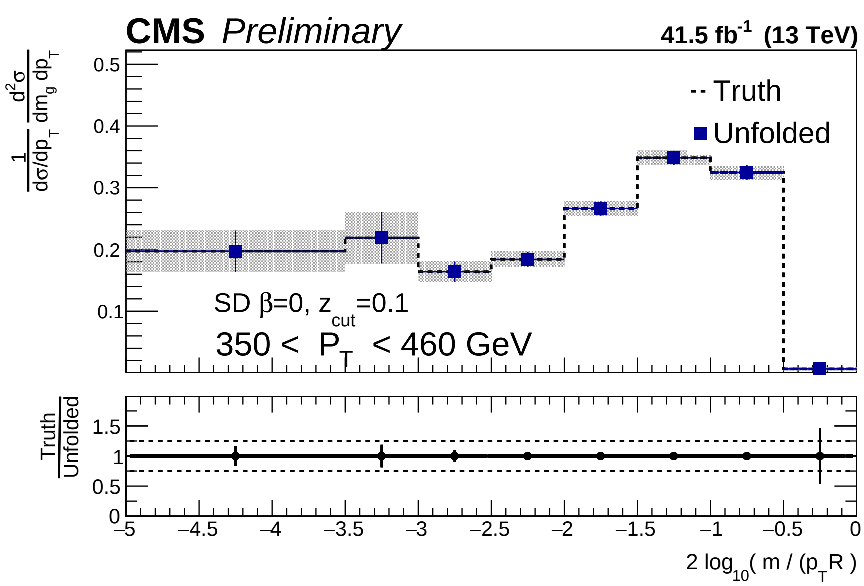

Figure 171: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

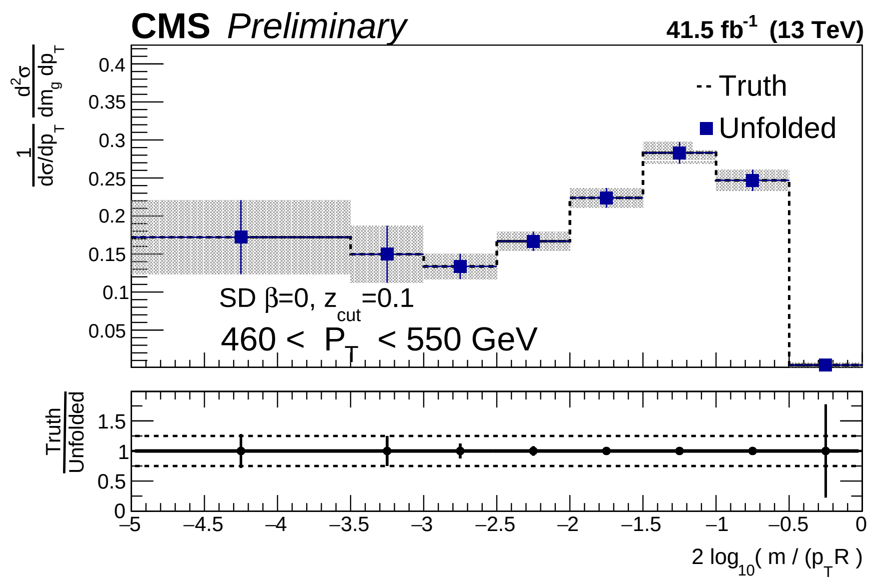

Figure 172: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_T 460-550 GeV.

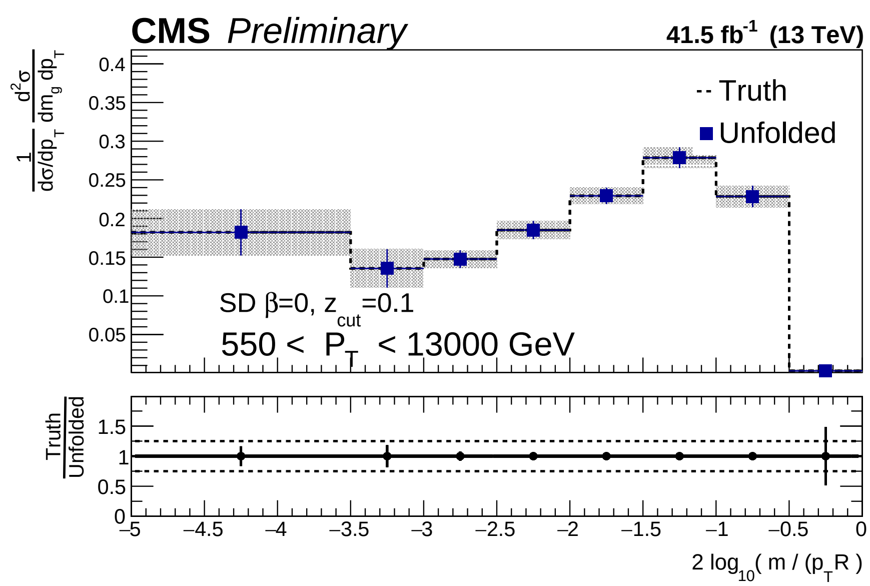

Figure 173: Dimensionless mass closure test of groomed reconstructed Monte Carlo, p_T 550-13000 GeV.

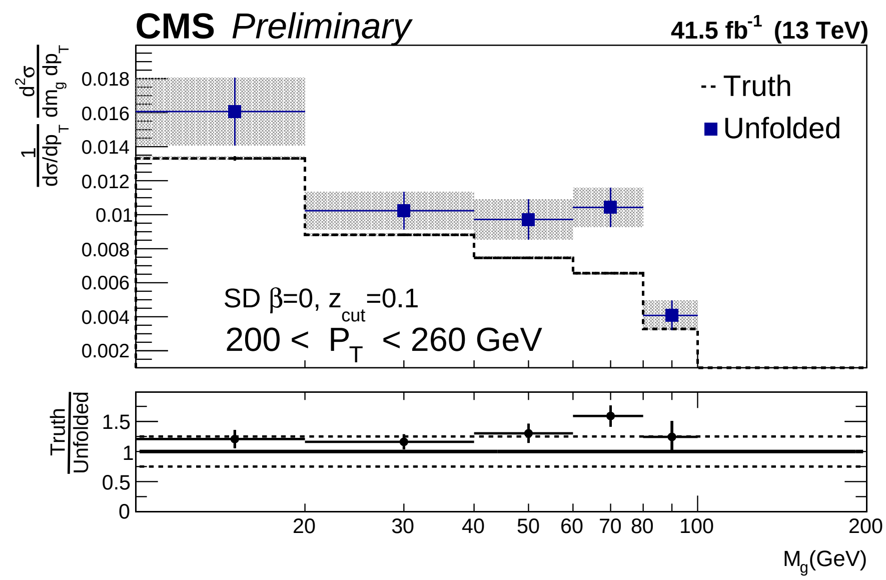

Figure 174: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

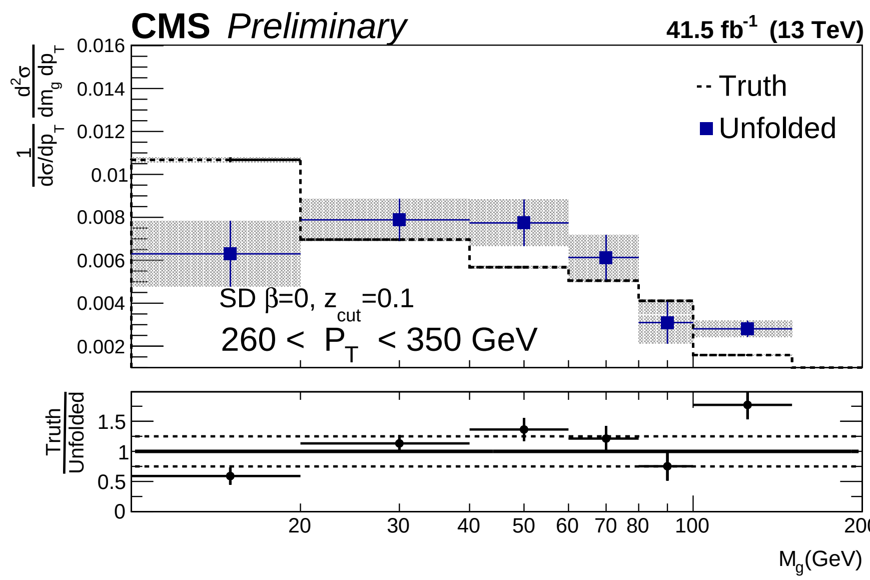

Figure 175: Physical bias test of groomed reconstructed Monte Carlo, p_T 260-350 GeV.

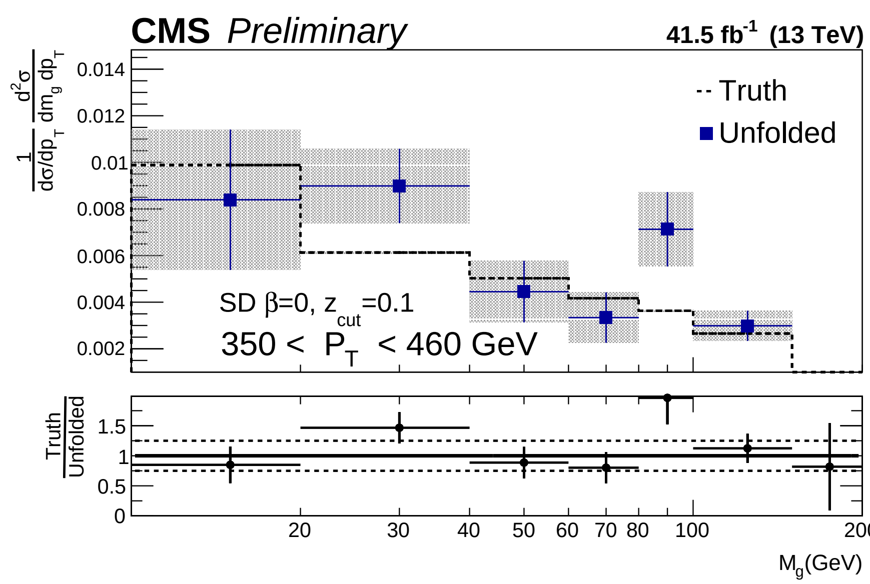

Figure 176: Physical bias test of groomed reconstructed Monte Carlo, p_T 350-460 GeV.

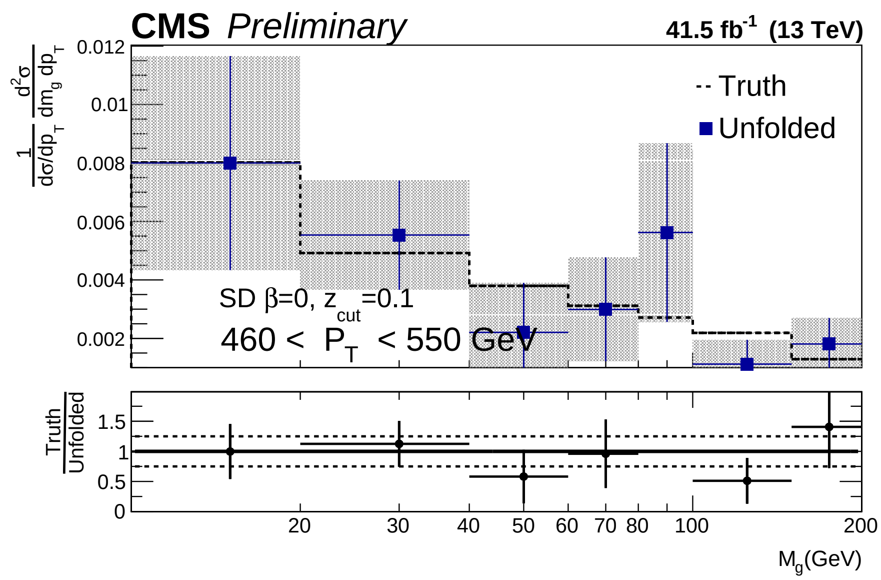

Figure 177: Physical bias test of groomed reconstructed Monte Carlo, p_T 460-550 GeV.

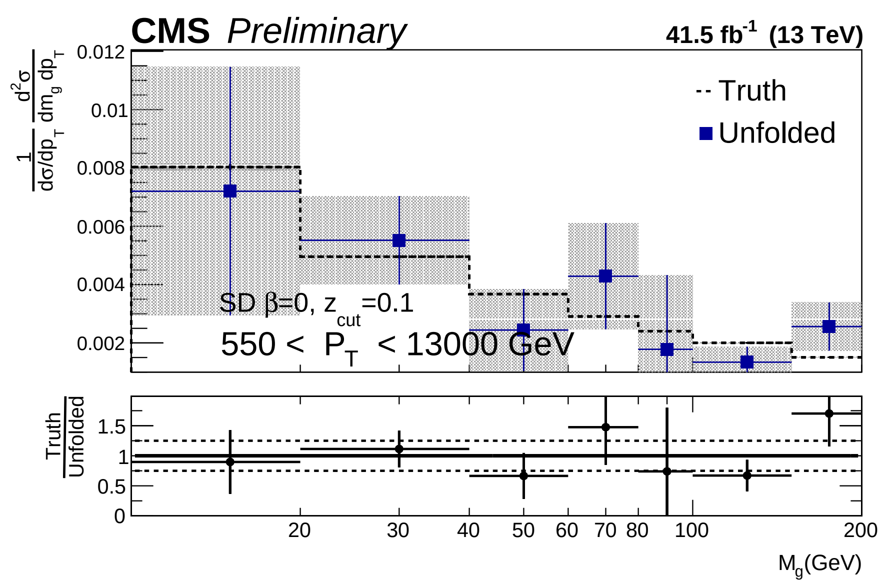

Figure 178: Physical bias test of groomed reconstructed Monte Carlo, p_T 550-13000 GeV.

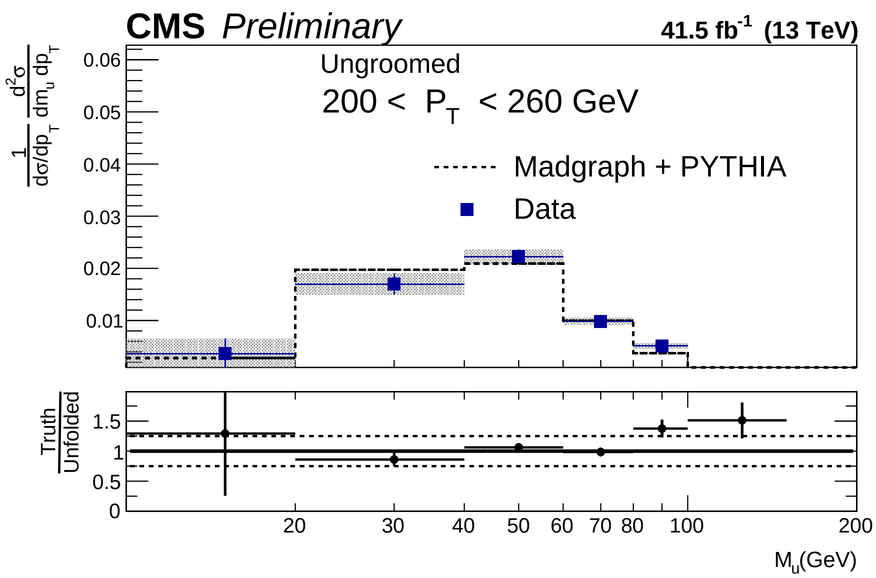

Figure 179: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 200-260 GeV.

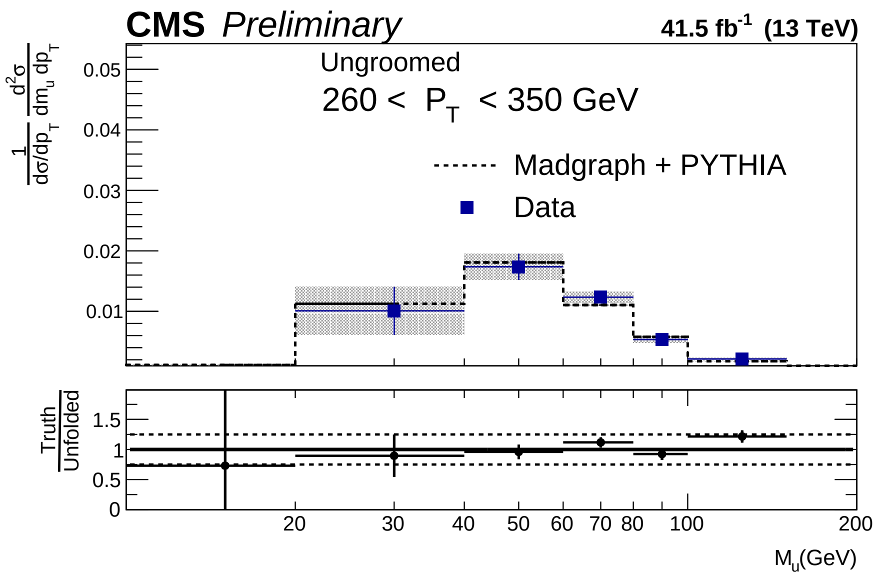

Figure 180: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 260-350 GeV.

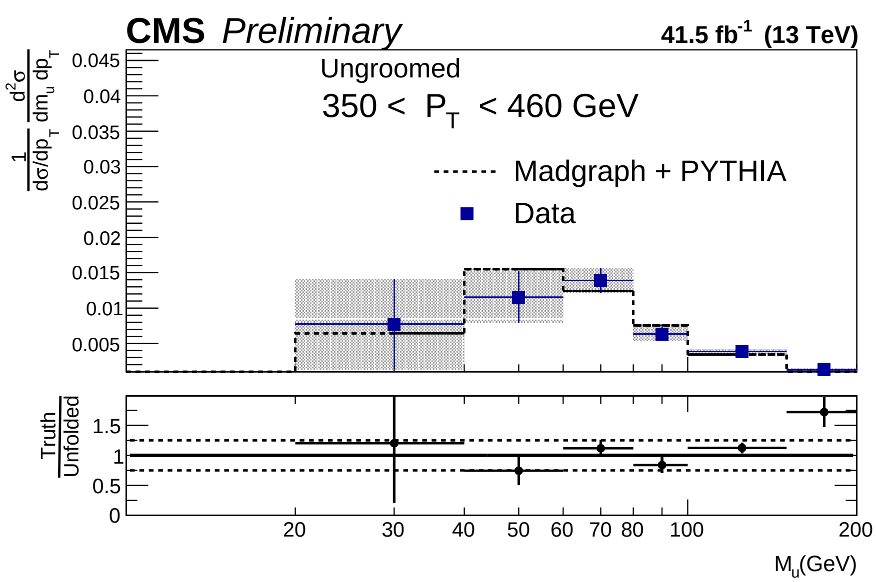

Figure 181: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 350-460 GeV.

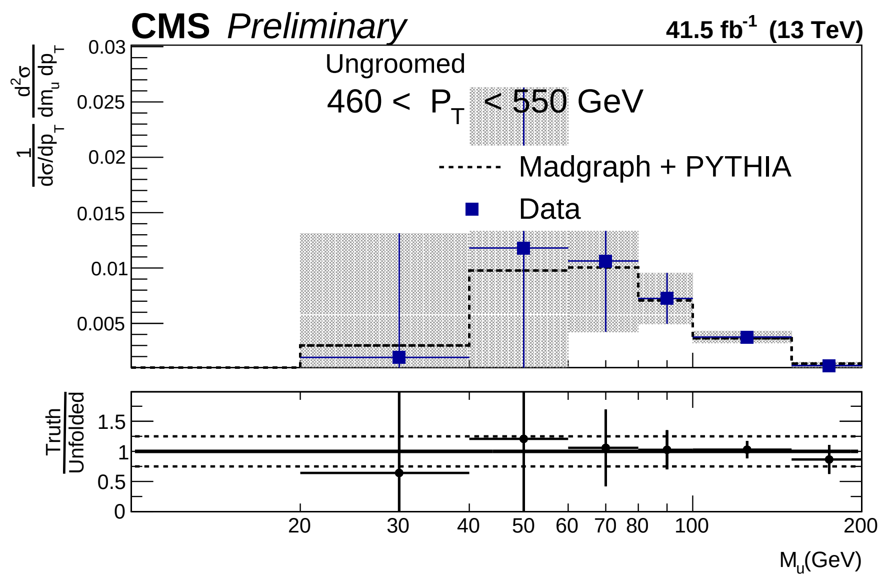

Figure 182: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 460-550 GeV.

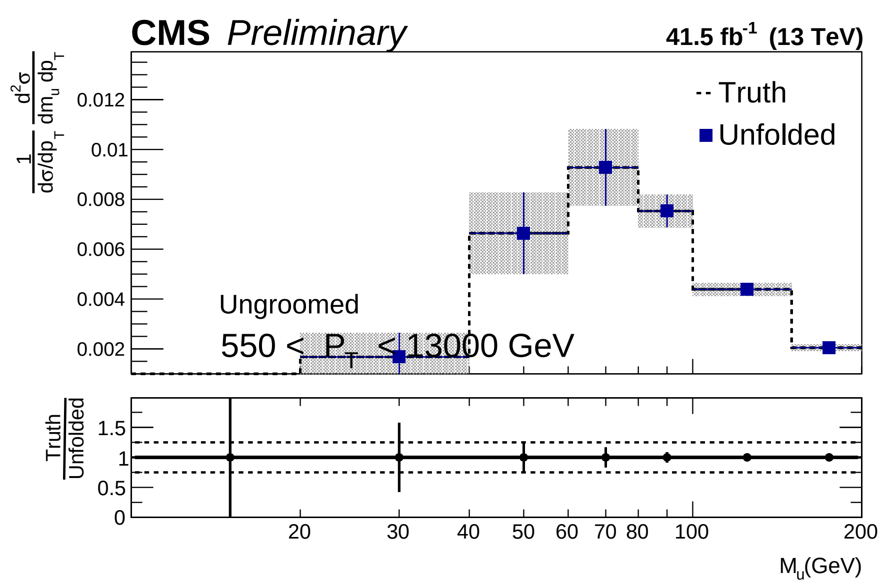

Figure 183: Normalized cross section results with respect to jet mass for ungroomed jets, p_T 550-13000 GeV.

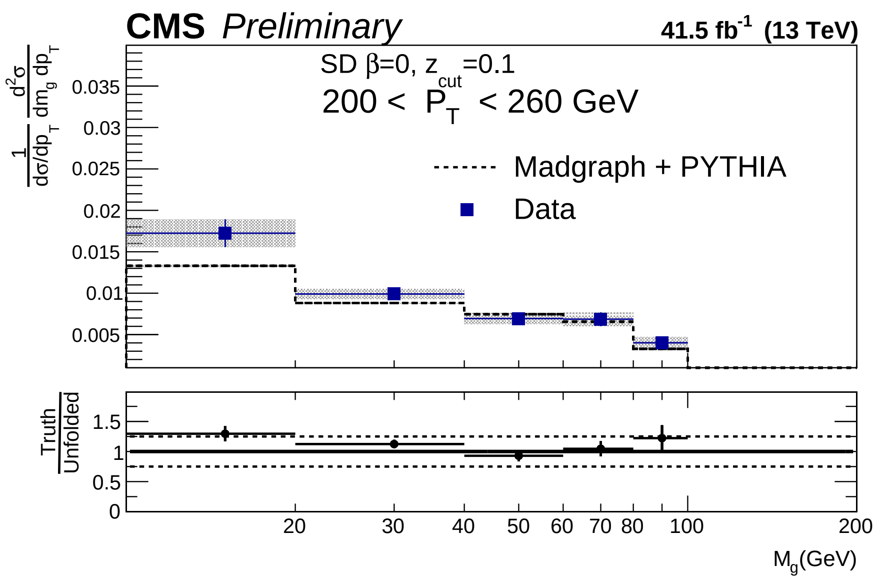

Figure 184: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 200-260 GeV.

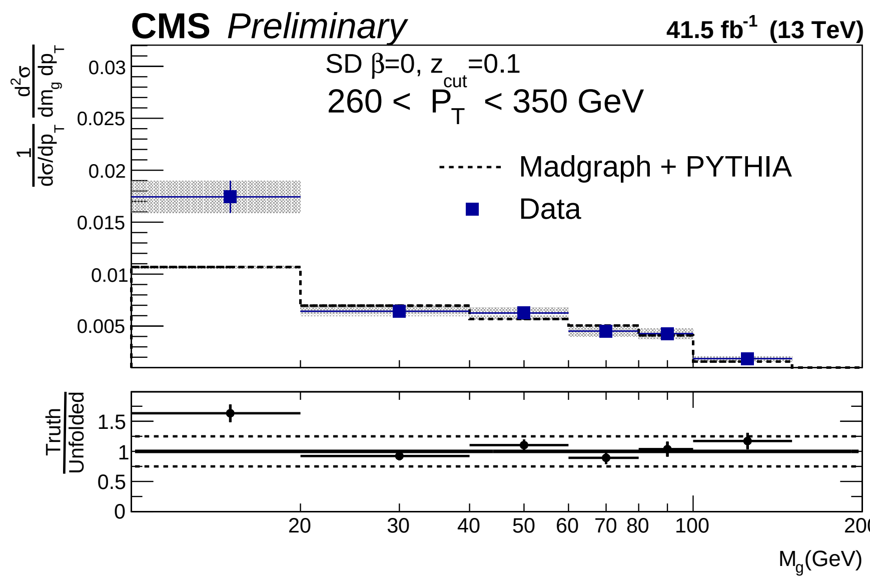

Figure 185: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 260-350 GeV.

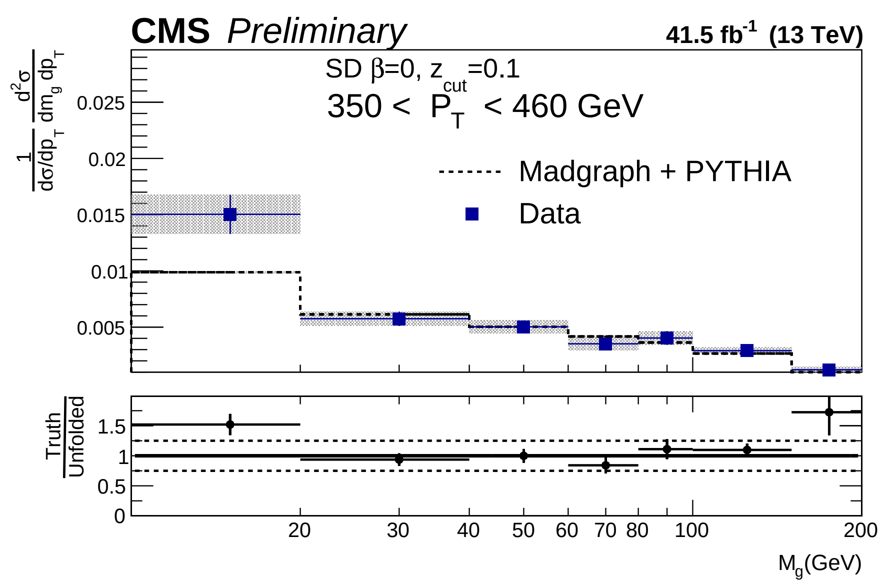

Figure 186: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 350-460 GeV.

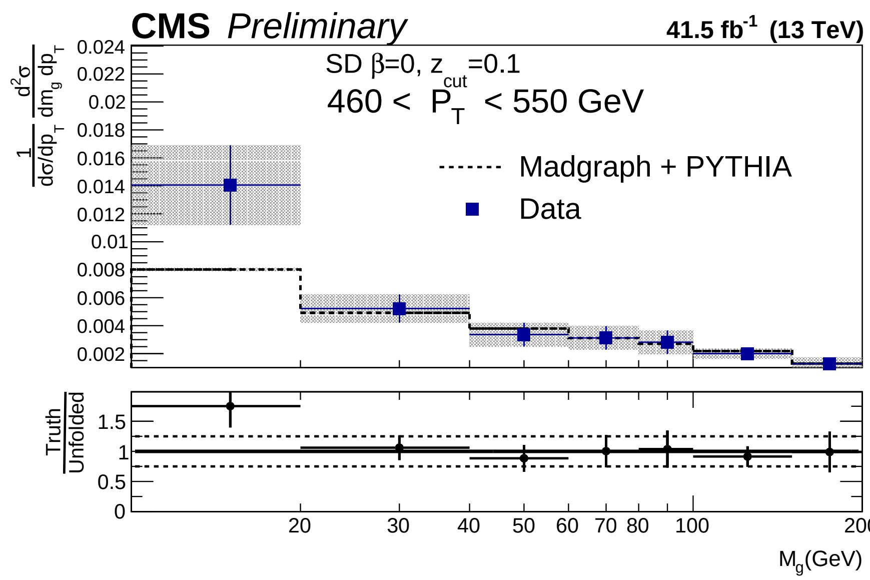

Figure 187: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 460-550 GeV.

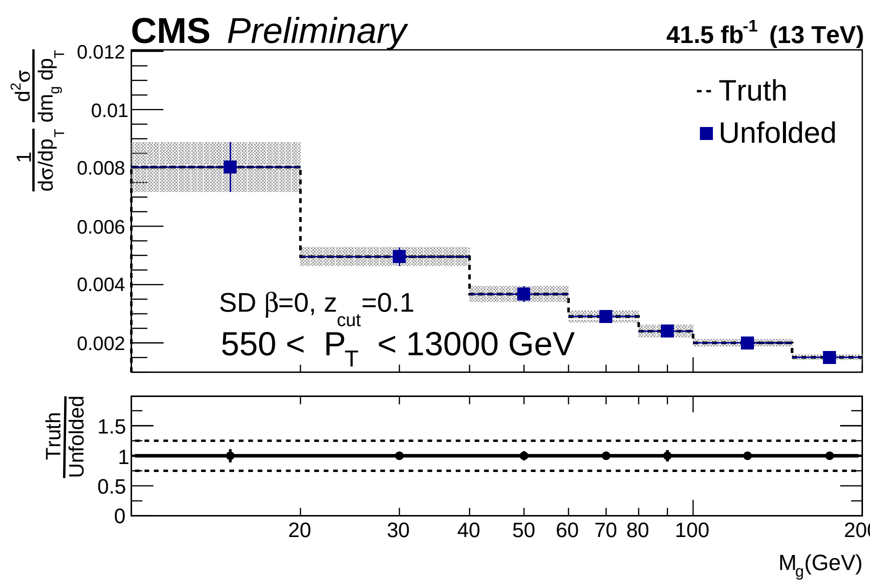

Figure 188: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 550-13000 GeV.

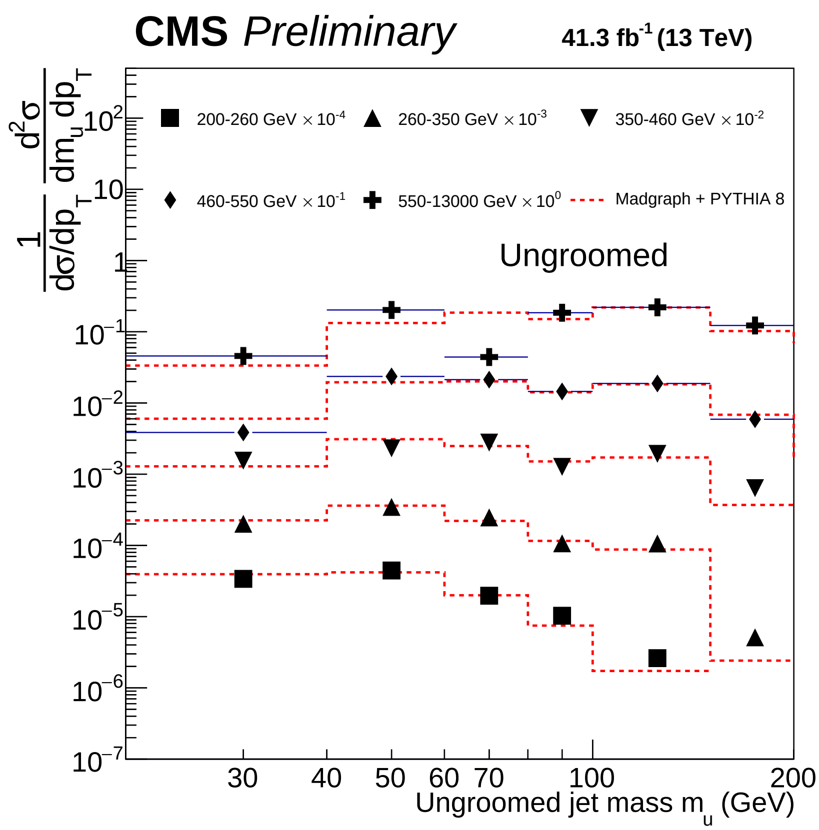

Figure 189: Results for ungroomed reconstructed unfolding with jet mass.

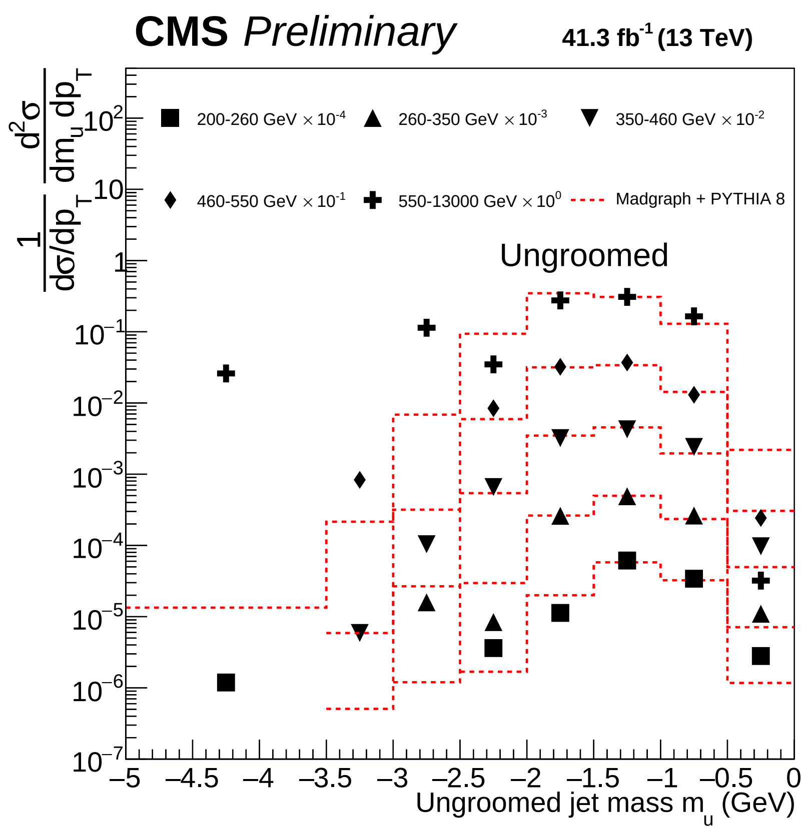

Figure 190: Results for ungroomed reconstructed unfolding with jet dimensionless mass.

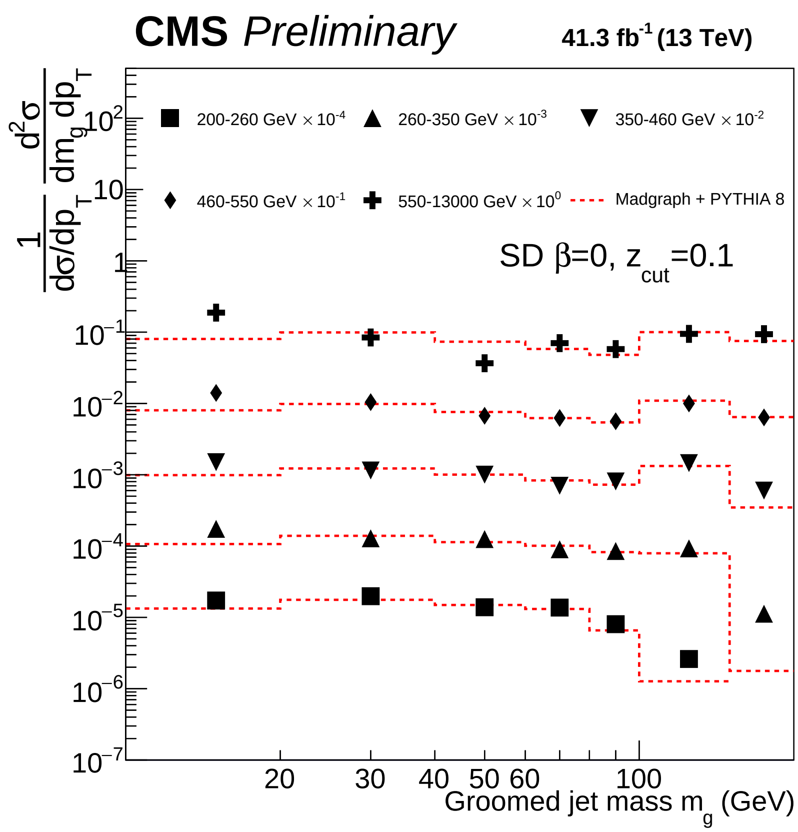

Figure 191: Results for groomed reconstructed unfolding with jet mass.

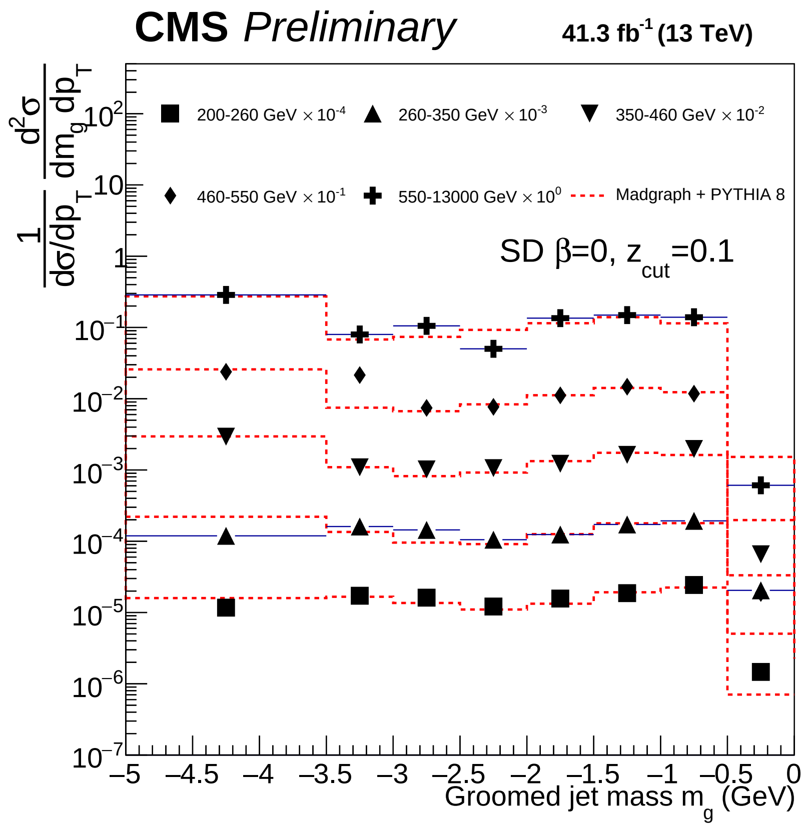

Figure 192: Results for groomed reconstructed unfolding with jet dimensionless mass.

The resulting correlation matrix without systematic uncertainties for the ungroomed jets is shown in Fig. 193, and for the groomed jets is shown in Fig. 195. The same figures with systematic uncertainties are shown for the ungroomed jets in Fig. 197, and for the groomed jets is shown in Fig. 199.

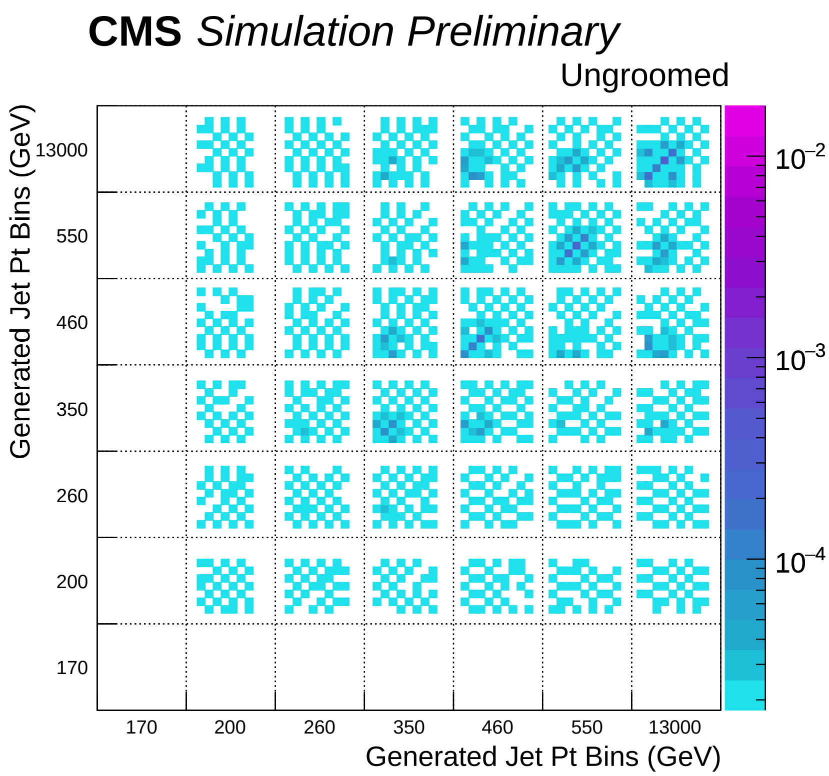

Figure 193: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,10,20,. . . 200 GeV), while the p_T is indexed by the major blocks (200,350,. . . 13000 GeV).

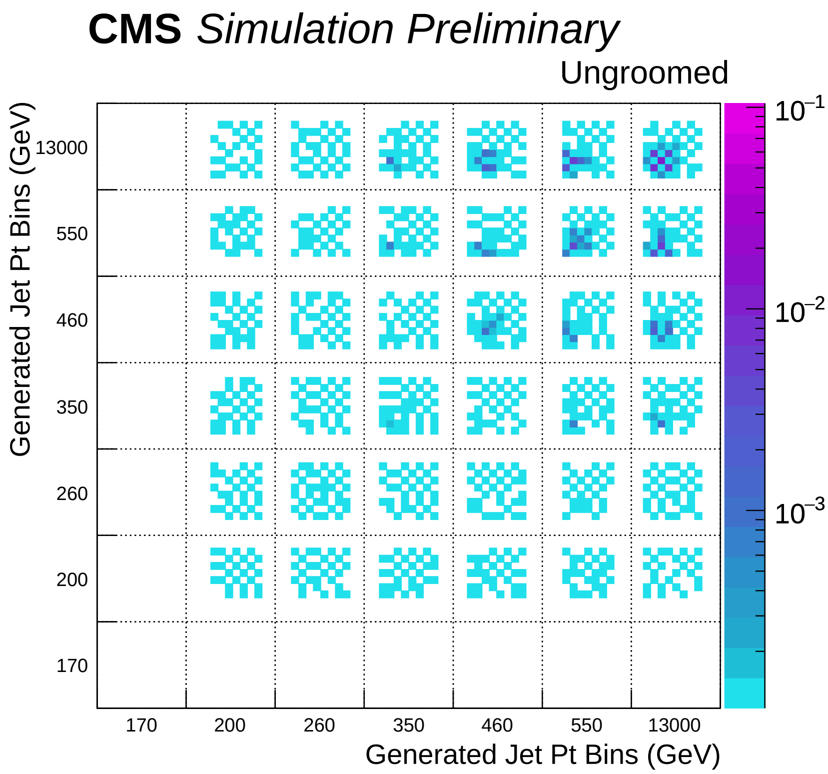

Figure 194: Correlation matrix for ungroomed jets. The bins are set such that the jet dimensionless mass is indexed by the minor blocks (-5, -3.5, -3. . . . 0), while the p_{T} is indexed by the major blocks (200,350,. . . 13000 GeV).

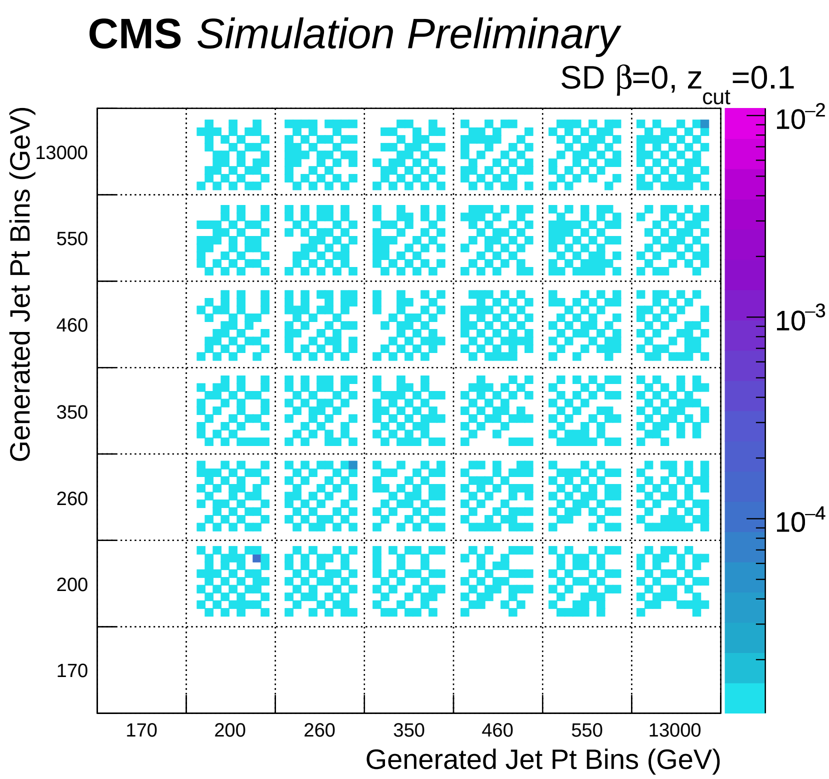

Figure 195: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 760 GeV).

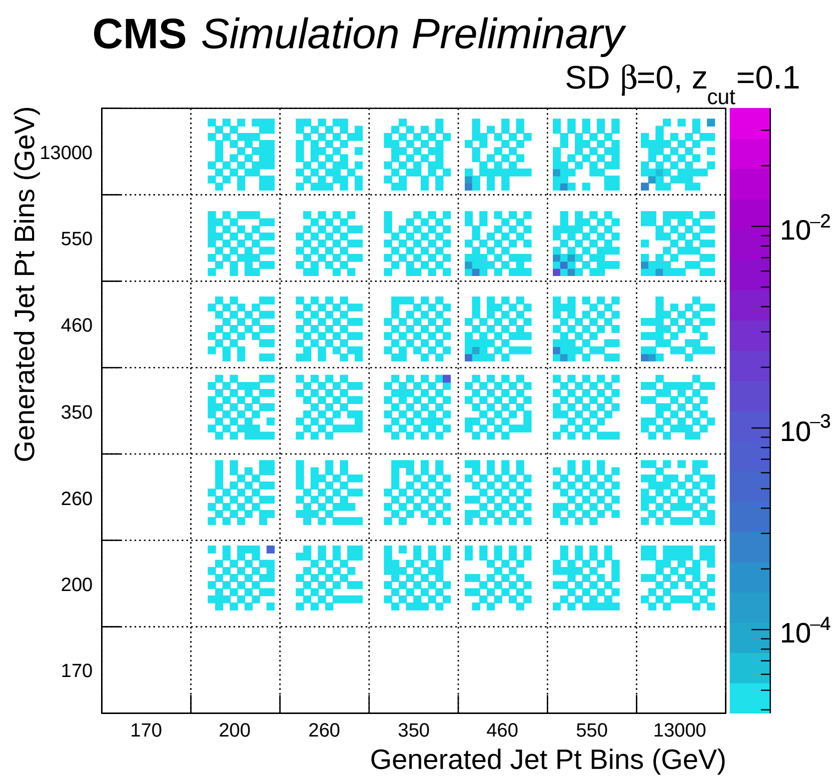

Figure 196: Correlation matrix for groomed jets. The bins are set such that the jet dimensionless mass is indexed by the minor blocks (-5, -3.5, -3. . . . 0), while the p_T is indexed by the major blocks (200,340,. . . 760 GeV).

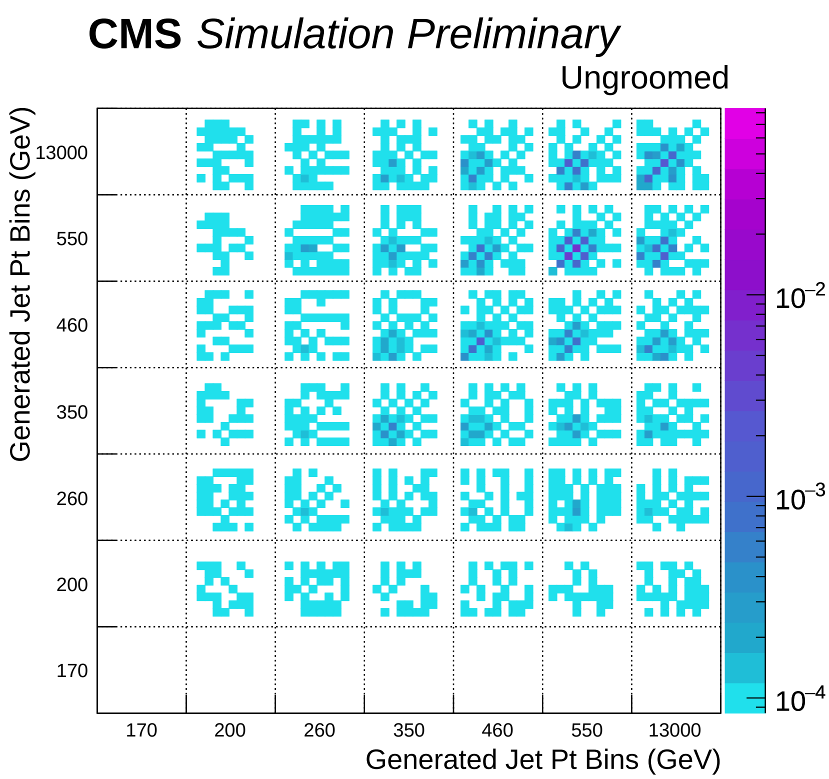

Figure 197: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,260,. . . 13000 GeV).

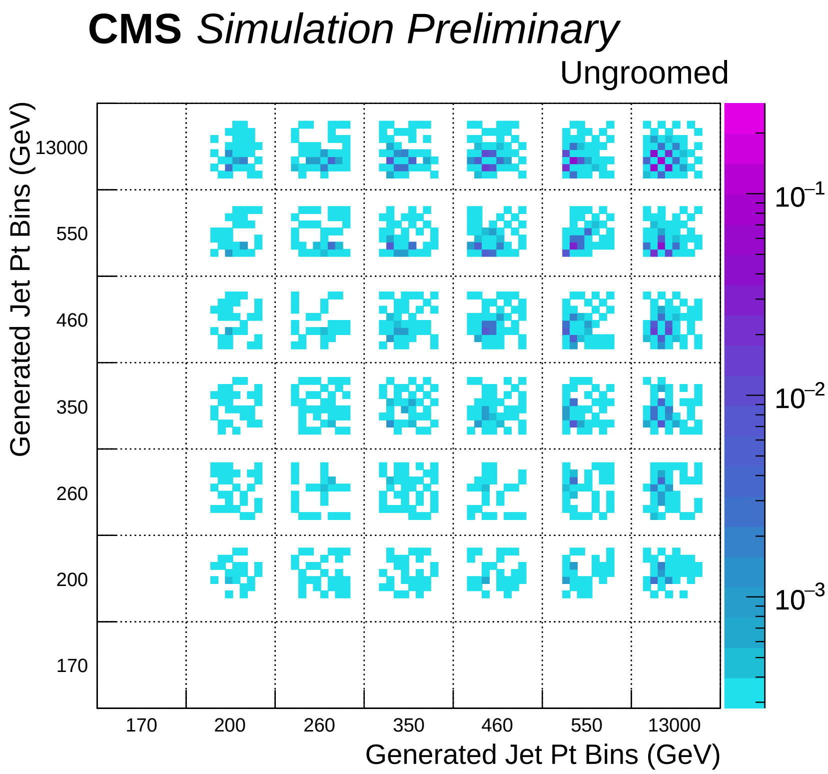

Figure 198: Correlation matrix for ungroomed jets. The bins are set such that the jet dimensionless mass is indexed by the minor blocks (-5, -3.5, -3. . . . 0), while the p_T is indexed by the major blocks (200,260,. . . 13000 GeV).

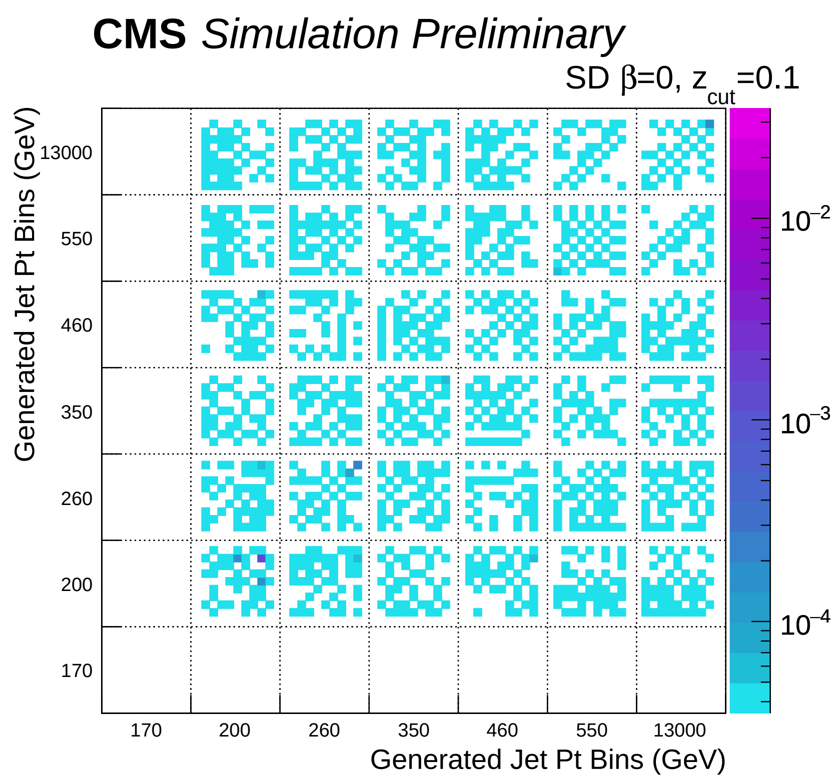

Figure 199: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_{T} is indexed by the major blocks (200,340,. . . 0 GeV).

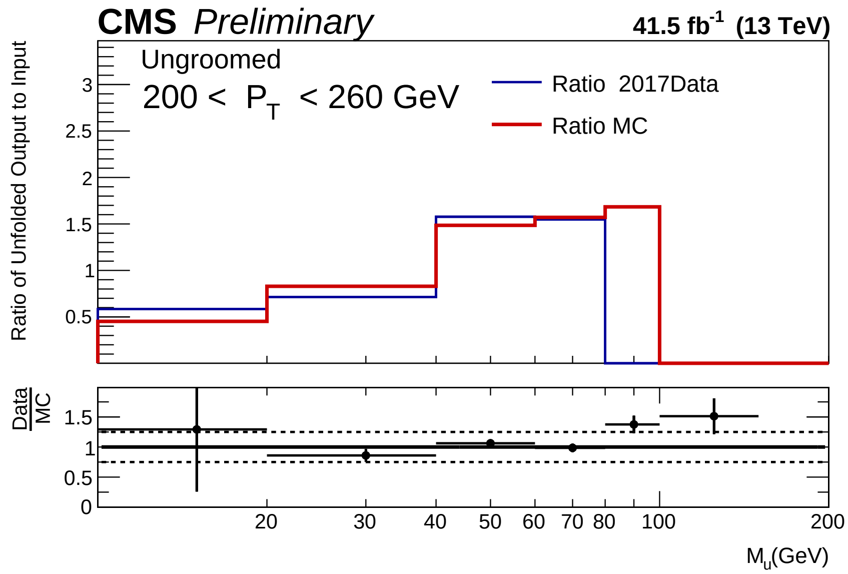

Figure 200: Ratio of unfolded over raw data and MC for ungroomed jets, p_{T} 200-260 GeV.

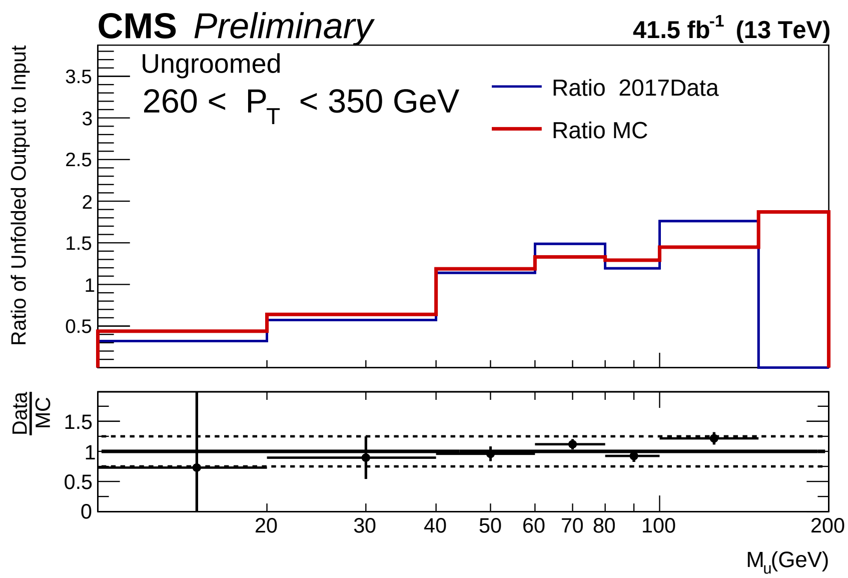

Figure 201: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 260-350 GeV.

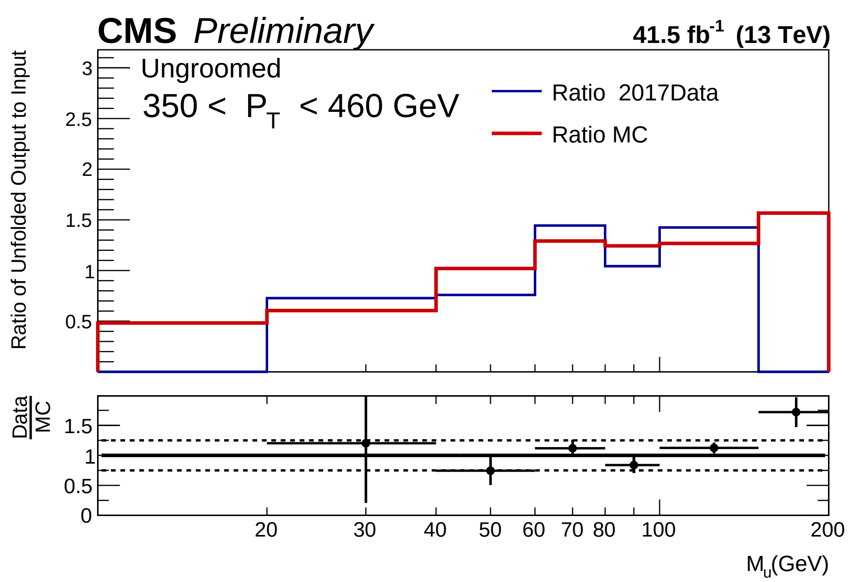

Figure 202: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 350-460 GeV.

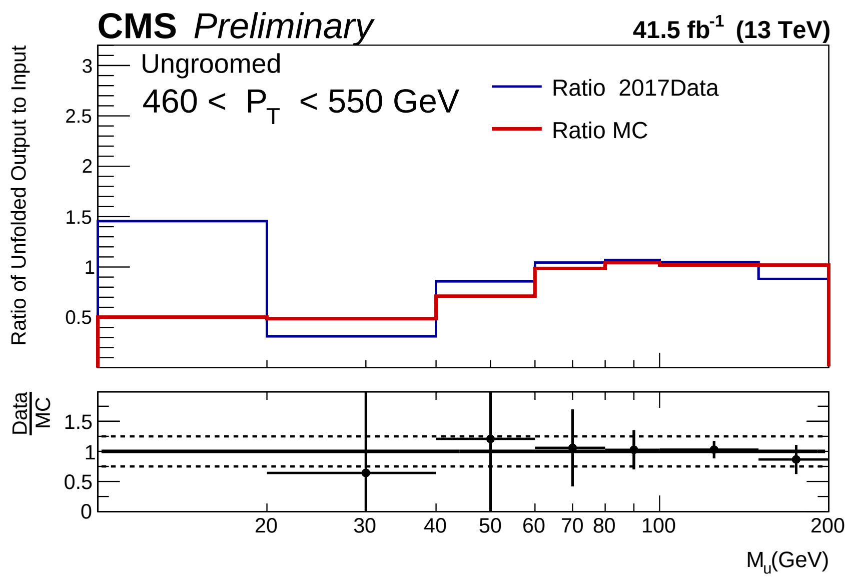

Figure 203: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 460-550 GeV.

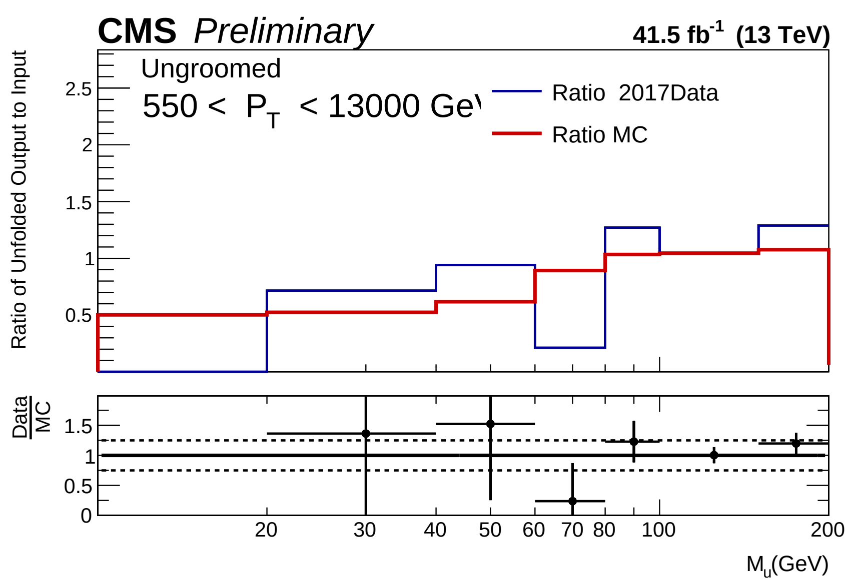

Figure 204: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 550-13000 GeV.

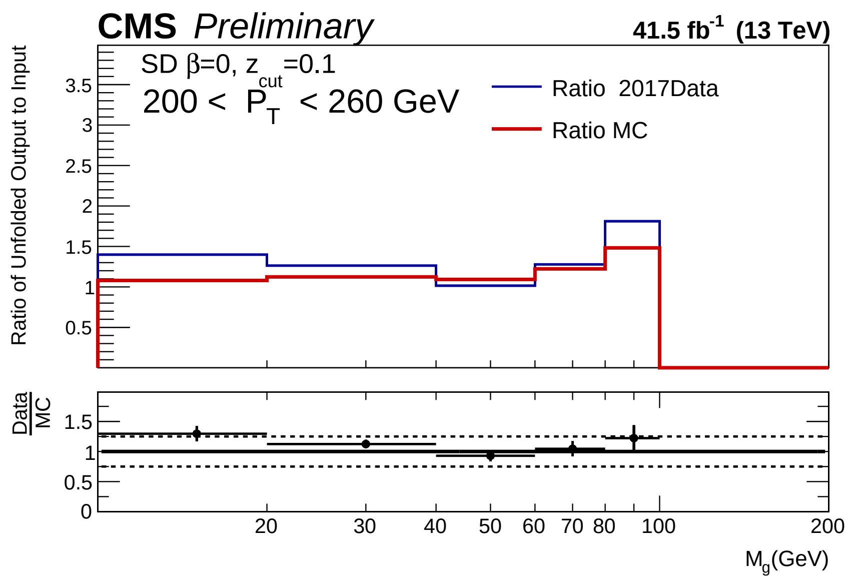

Figure 205: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 200-260 GeV.

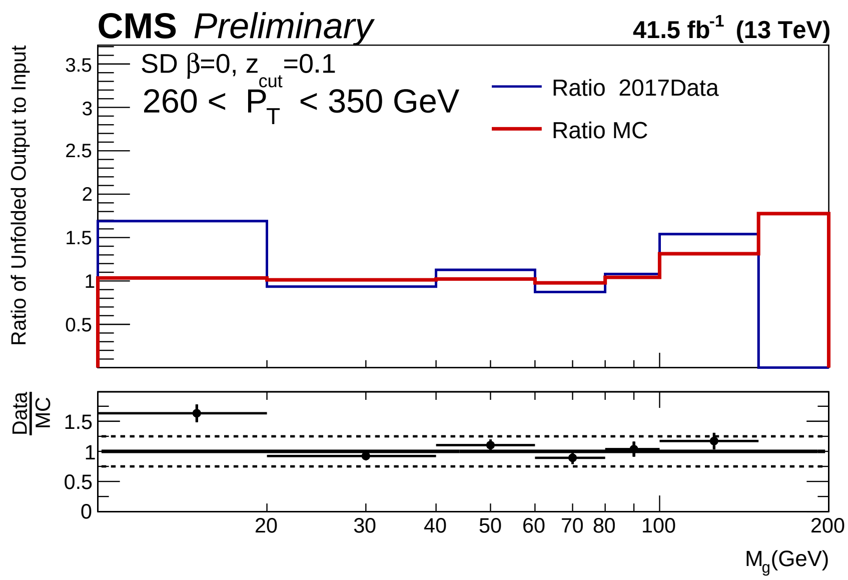

Figure 206: Ratio of unfolded over raw data and MC for groomed jets, p_T 260-350 GeV.

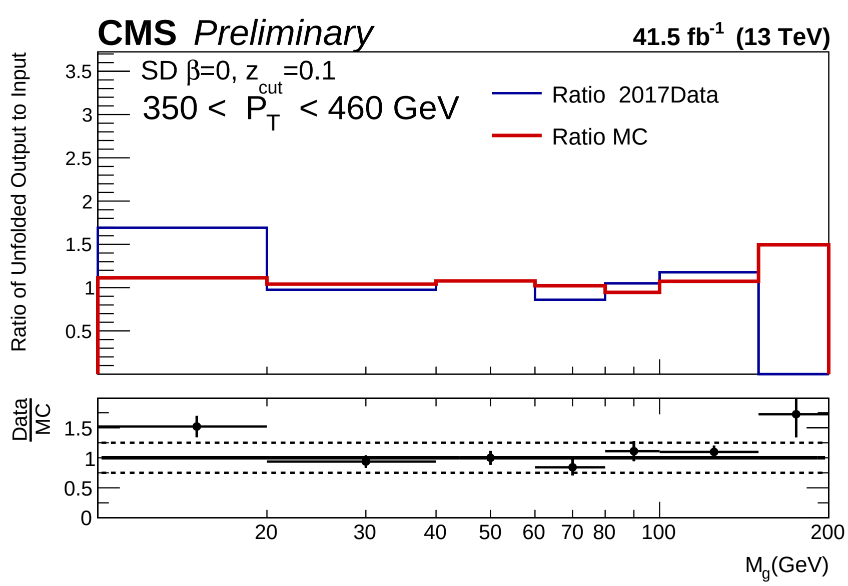

Figure 207: Ratio of unfolded over raw data and MC for groomed jets, p_T 350-460 GeV.

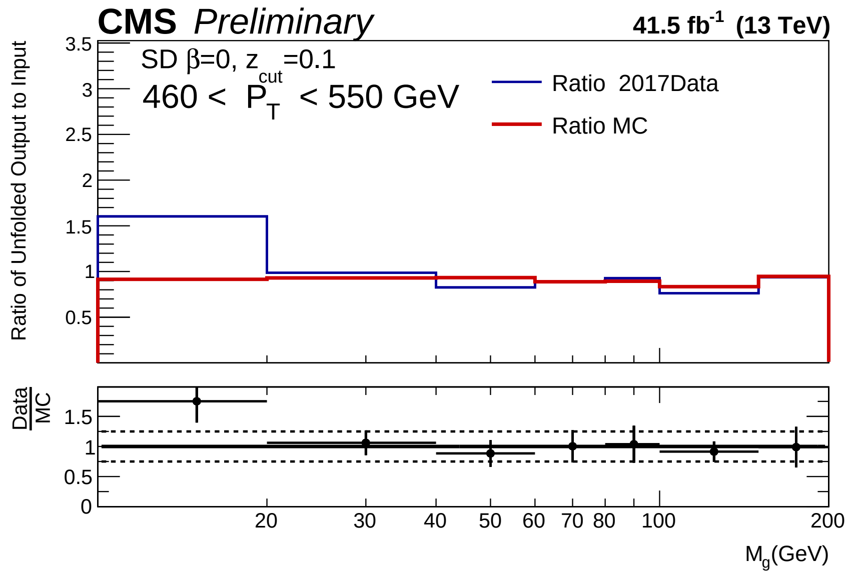

Figure 208: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 460-550 GeV.

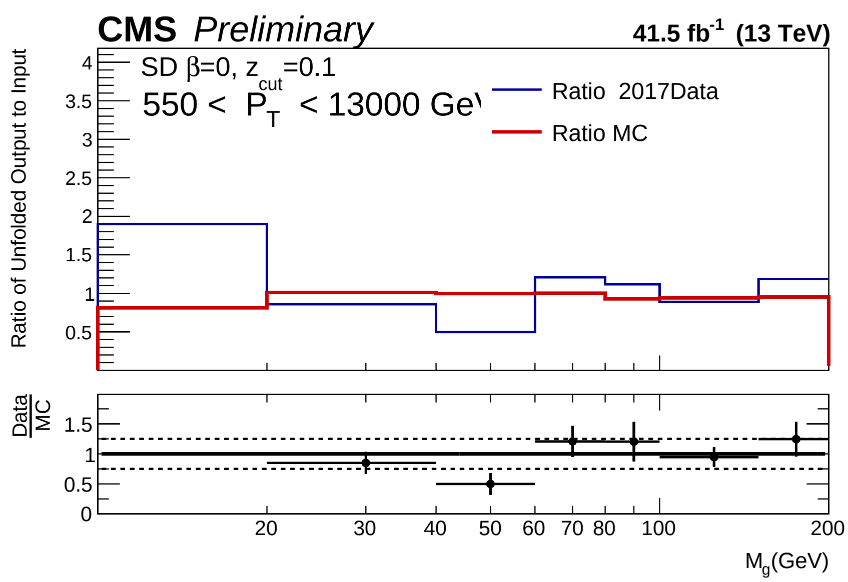

Figure 209: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 550-13000 GeV.

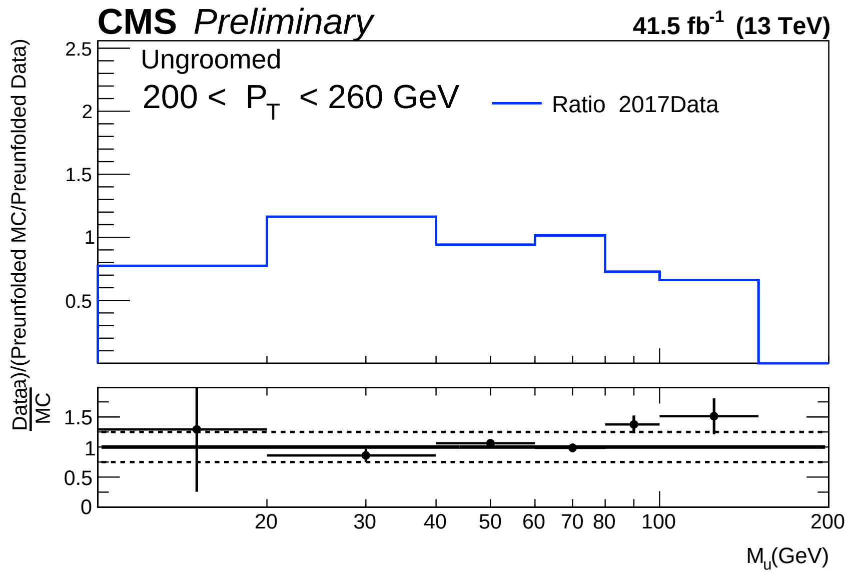

Figure 210: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 200-260 GeV.

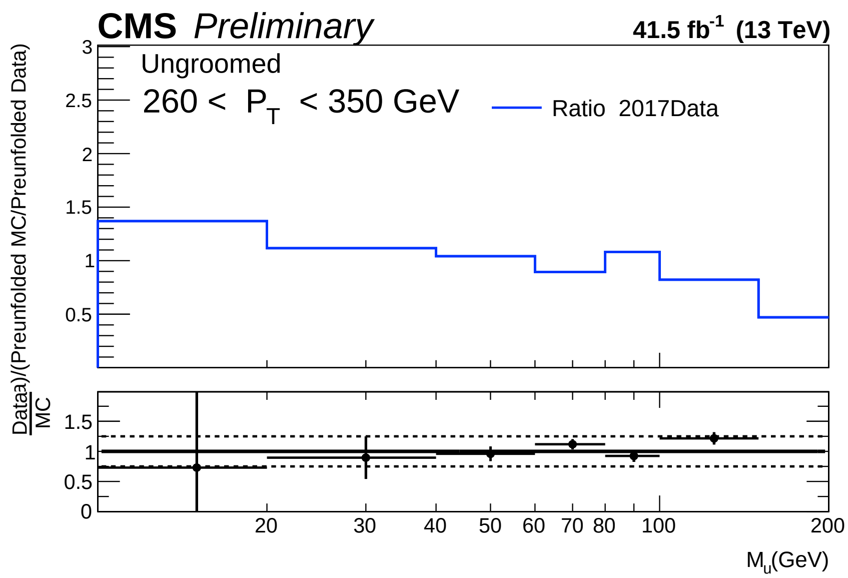

Figure 211: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 260-350 GeV.

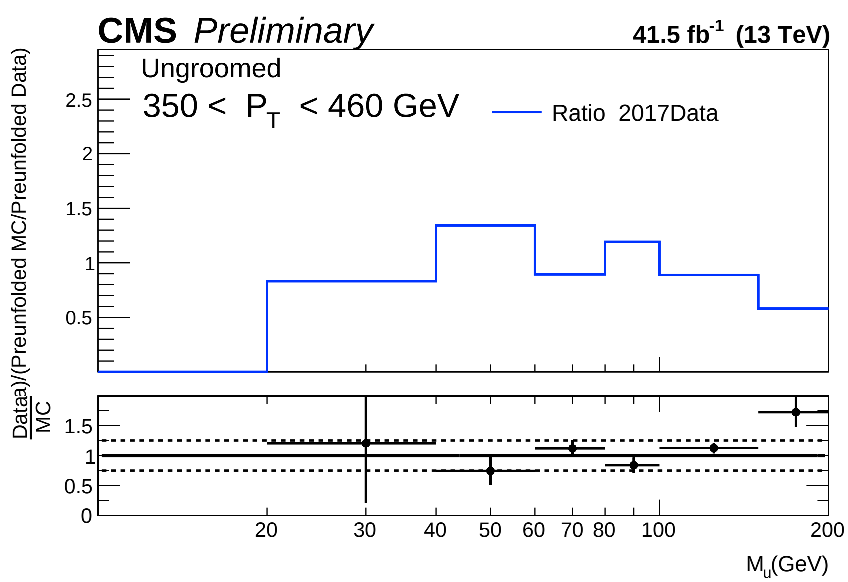

Figure 212: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 350-460 GeV.

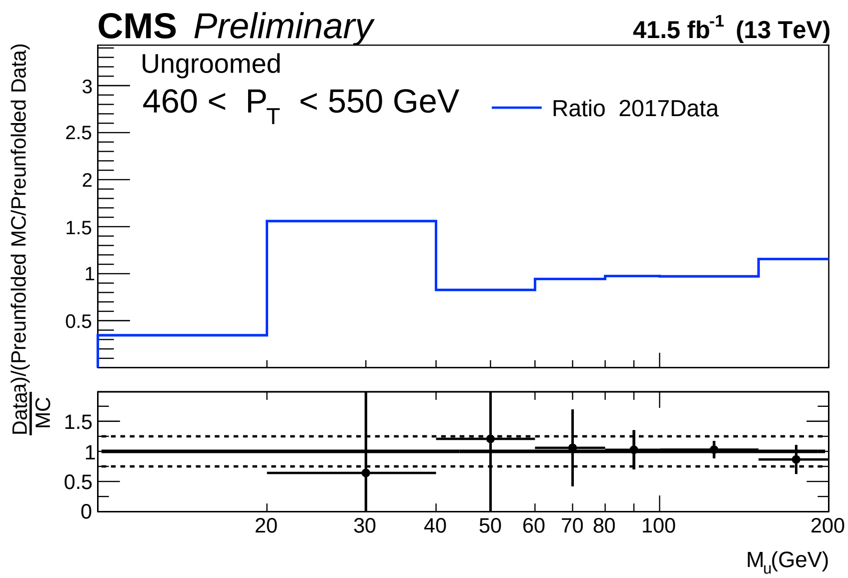

Figure 213: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 460-550 GeV.

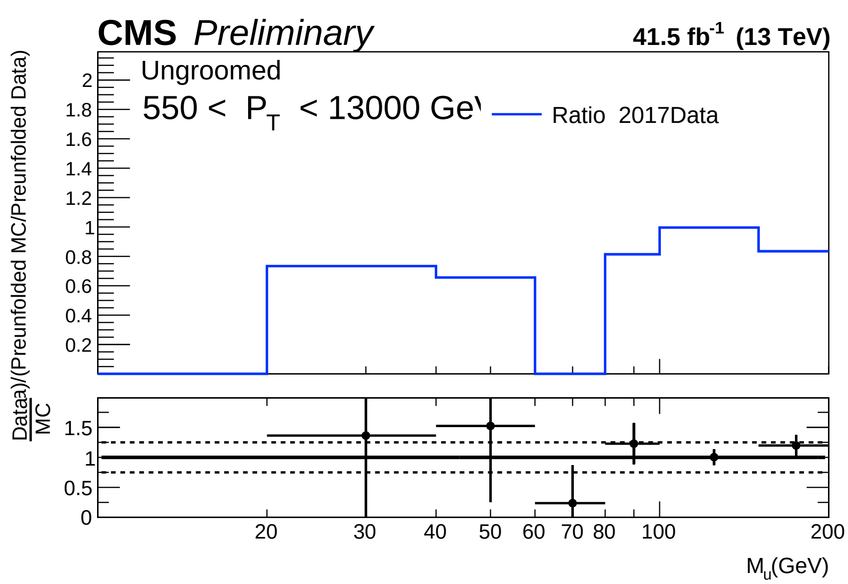

Figure 214: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 550-13000 GeV.

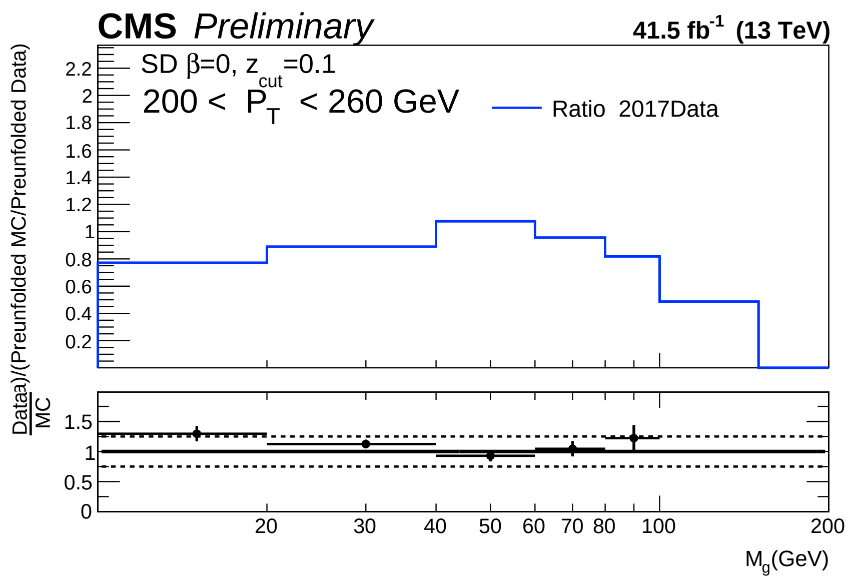

Figure 215: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 200-260 GeV.

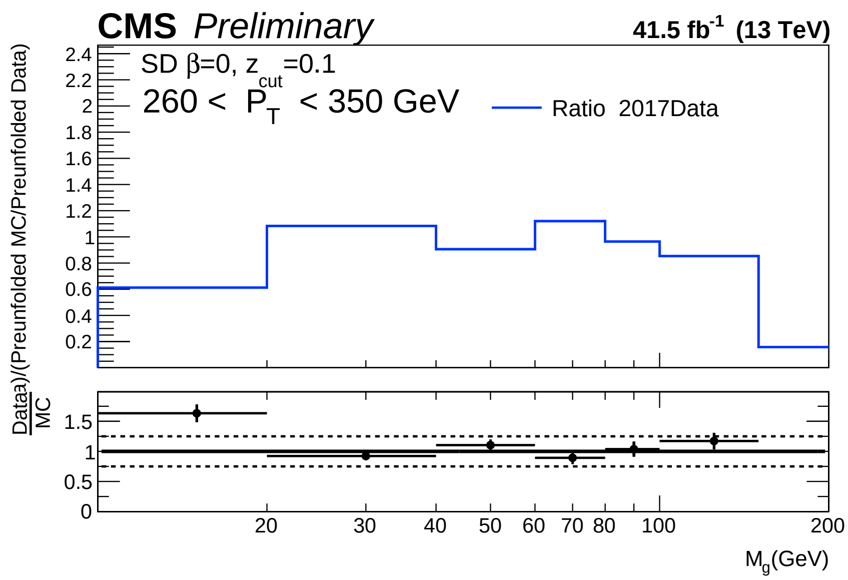

Figure 216: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 260-350 GeV.

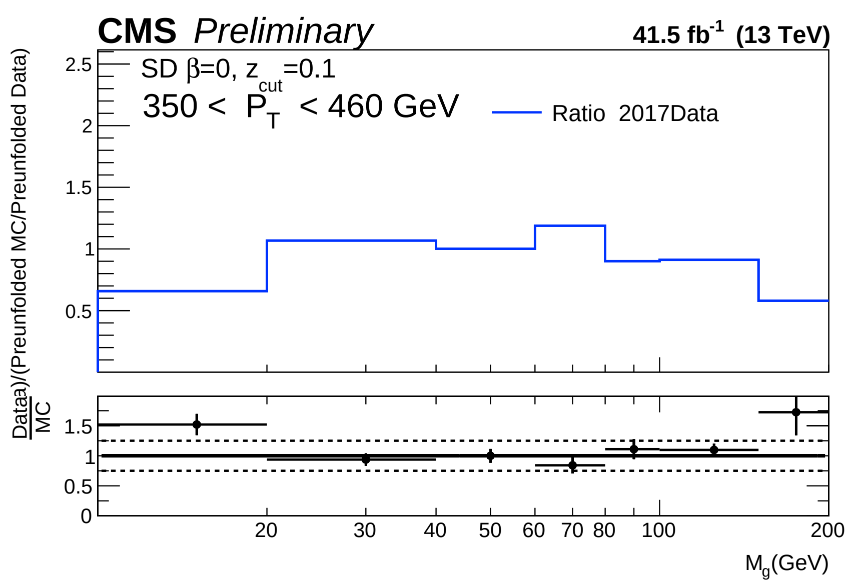

Figure 217: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 350-460 GeV.

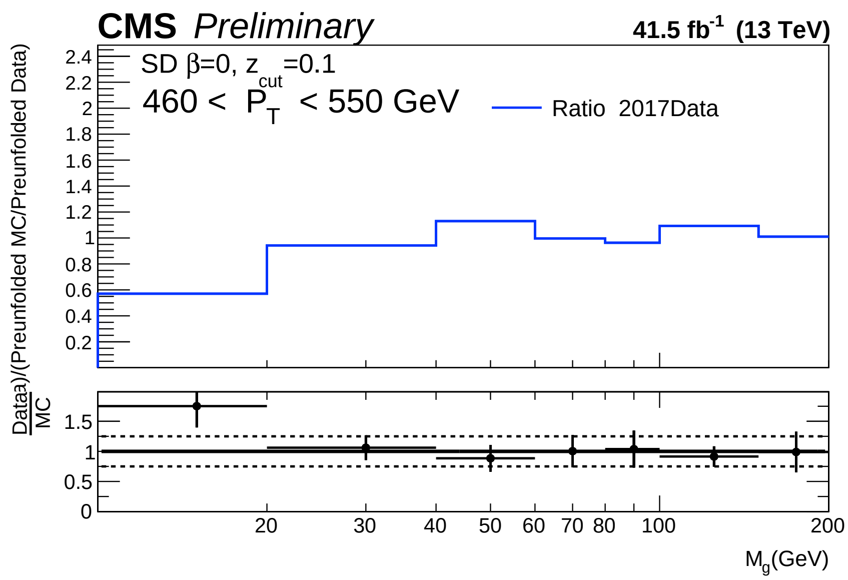

Figure 218: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_T 460-550 GeV.

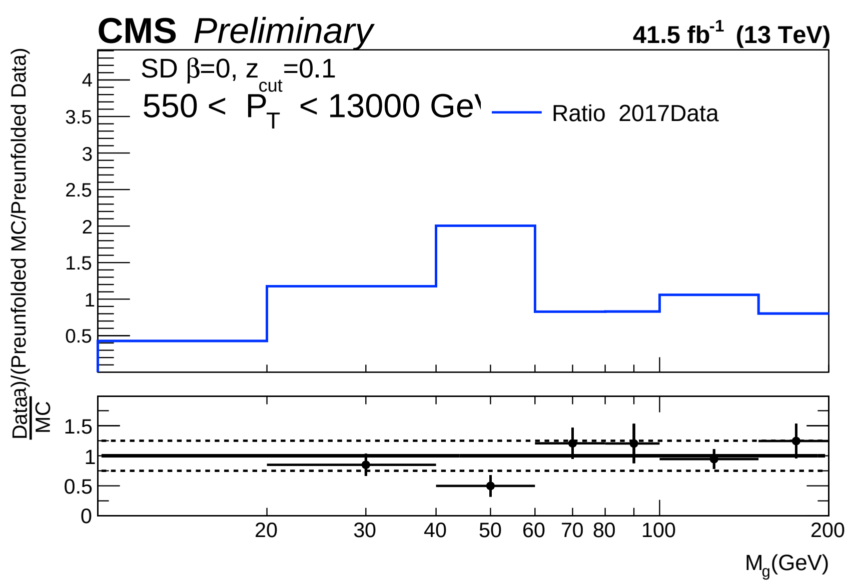

Figure 219: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_T 550-13000 GeV.

.3 2018 data results

This Appendix shows the distributions from the ”Detector Response” through the ”Results” sections of the main analysis note with only the 2018 data rather than the full Run 2 statistics seen in the main body of the note.

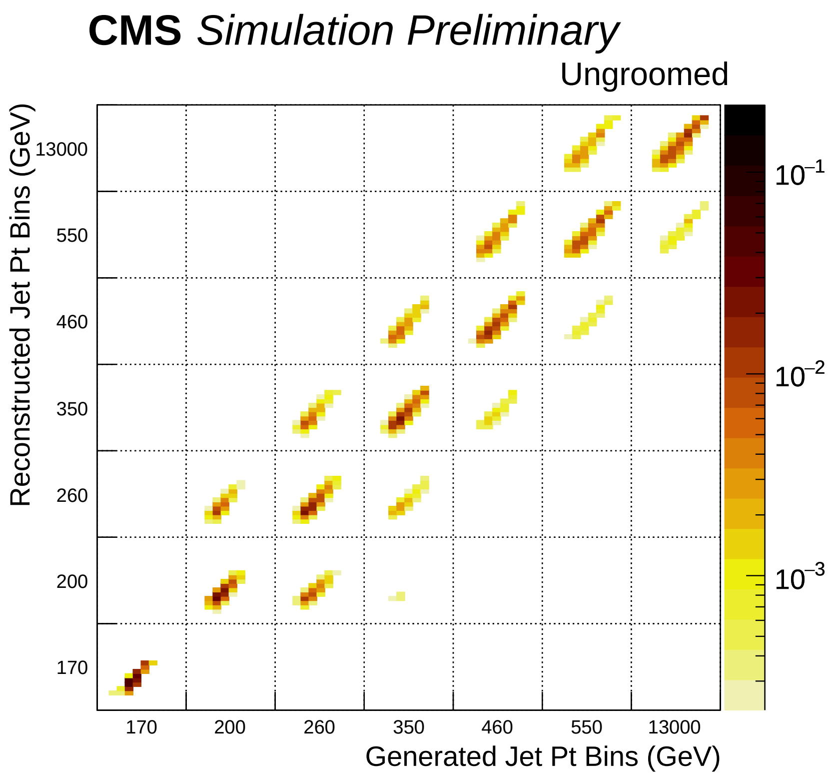

Figure 220: Two-dimensional response matrix for ungroomed jets. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV). While the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

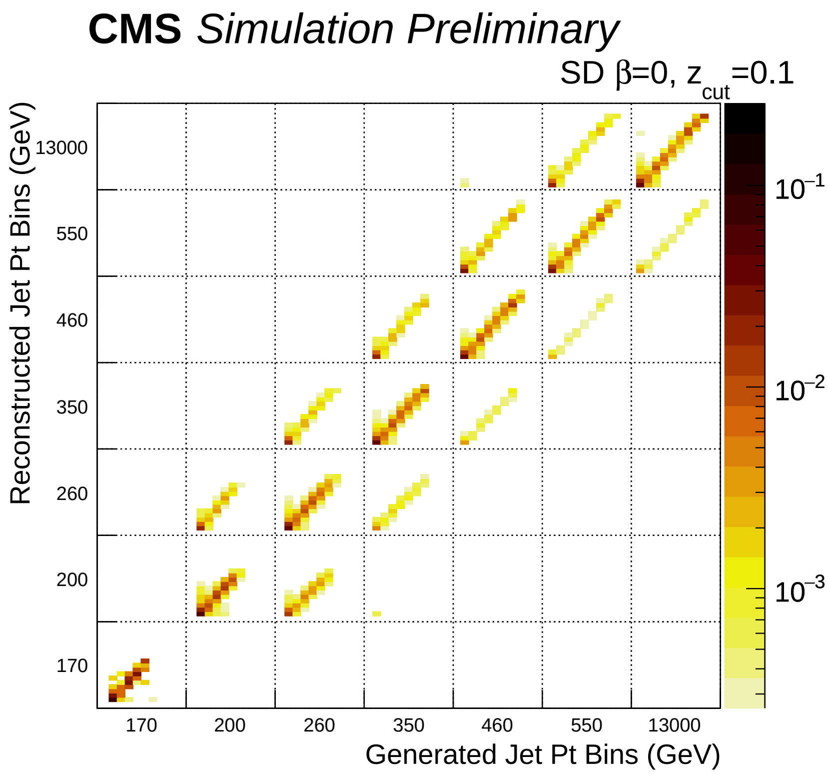

Figure 221: Two-dimensional response matrix for groomed jets $\beta = 0$, $z_{cut} = 0.1$. The bins are set such that the p_{T} is indexed by the major blocks (200, 260, 350, 460, 550, 13000 GeV).while the jet mass is indexed by the minor blocks on the x (y) axis 0, 10, 20, 40, 60, 80, 100, 150, 200, 13000 (0, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 200, 1000, 13000) GeV.

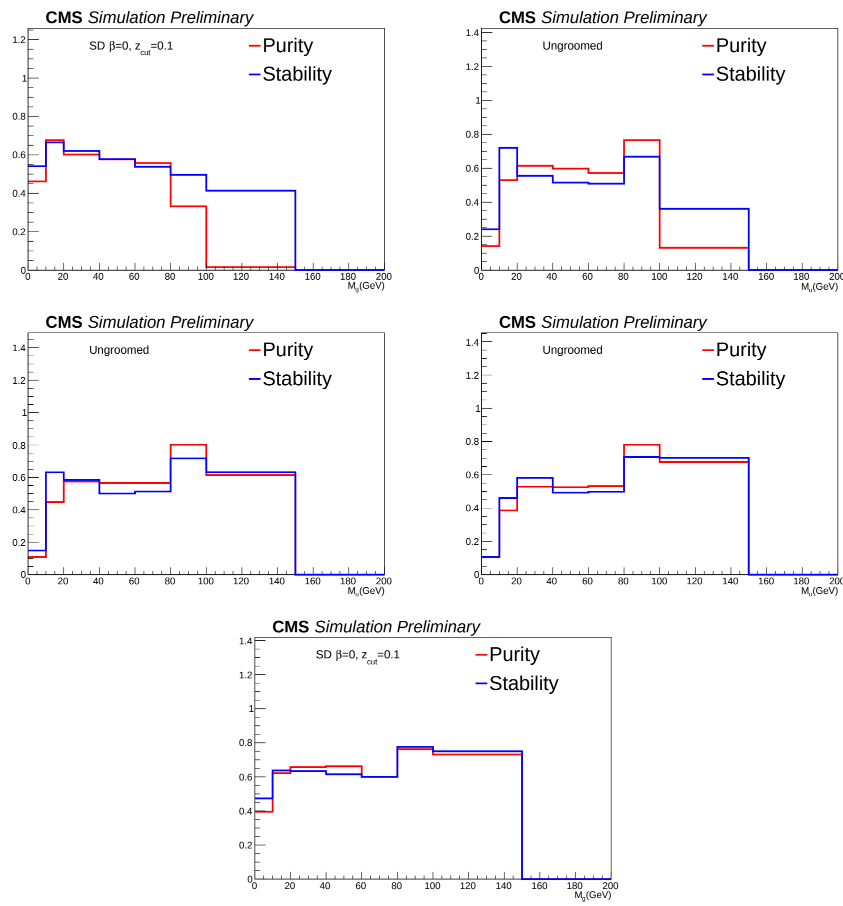

Figure 222: Purity and stability for ungroomed jets in various p_{T} bins.

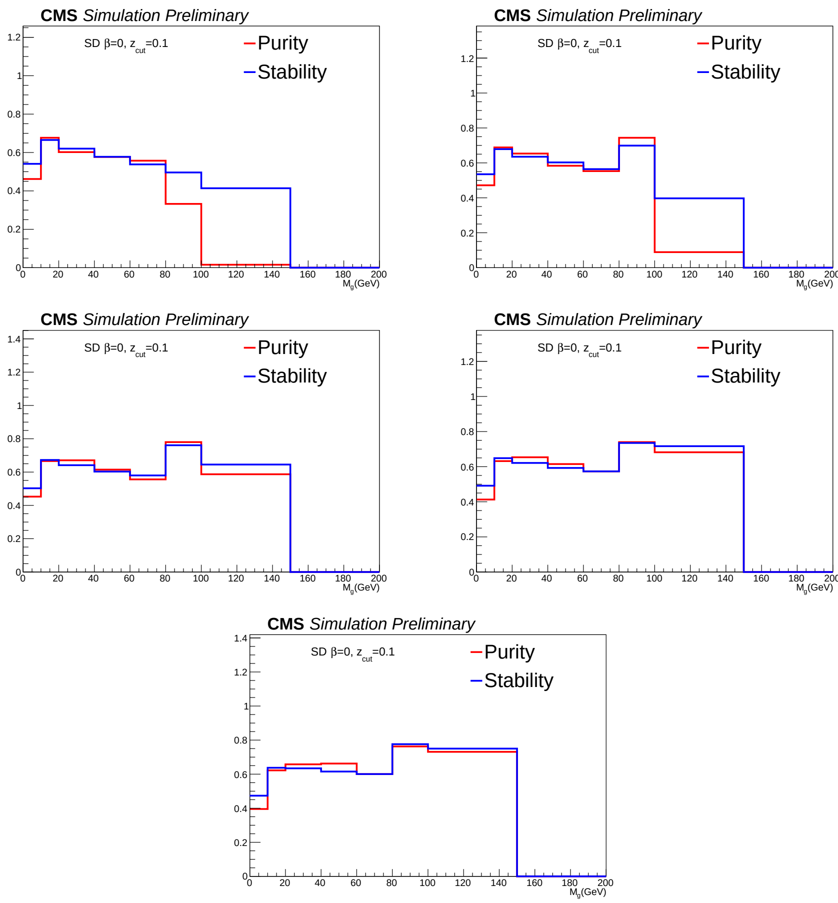

Figure 223: Purity and stability for groomed jets, where the soft-drop criterion was applied with $\beta = 0$ and $z_{cut} = 0.1$, in various p_{T} bins.

Figure-224 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_T bin for the normalized cross section.

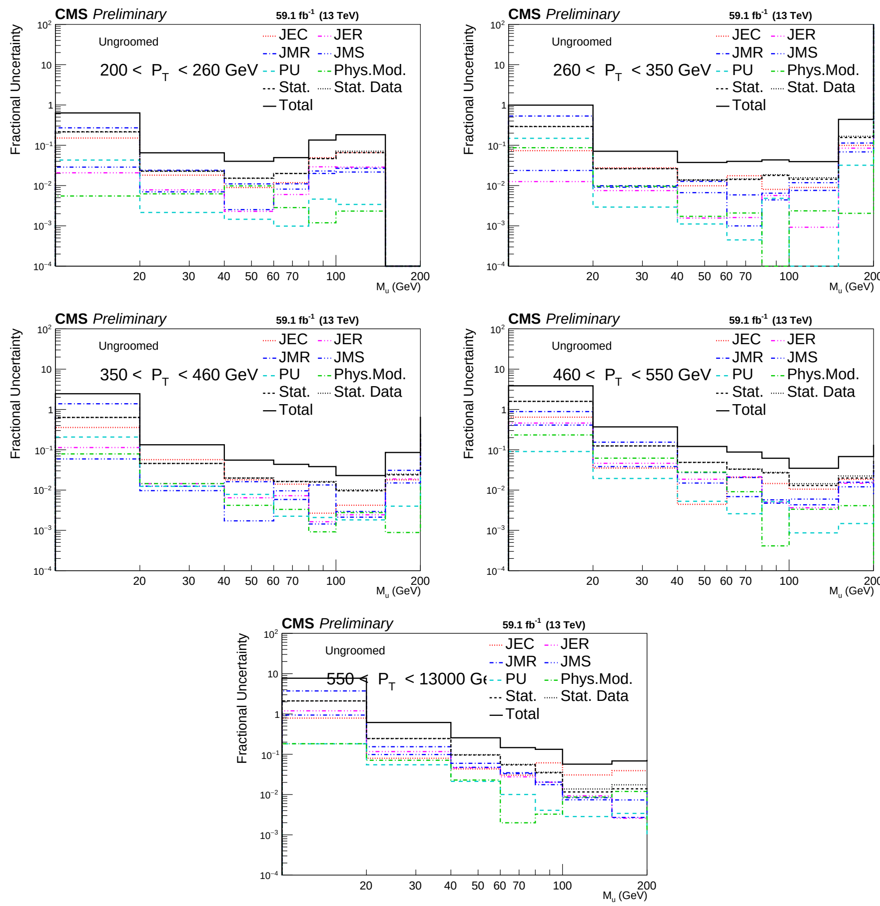

Figure 224: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the raw jet mass.

Figure-225 shows the uncertainty bands for all of the various sources listed above for each jet mass and p_T bin for the normalized cross section.

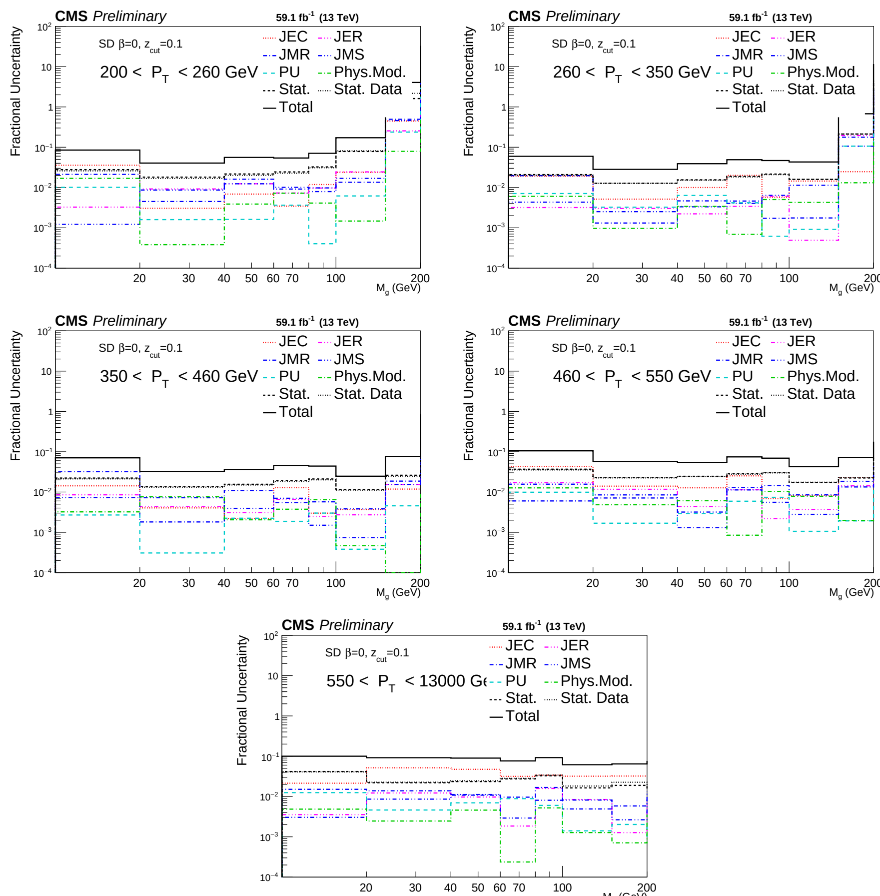

Figure 225: Systematic uncertainties are significantly reduced in the normalized cross section after unfolding for the various p_T bins for the jet mass after grooming with default Soft-Drop.

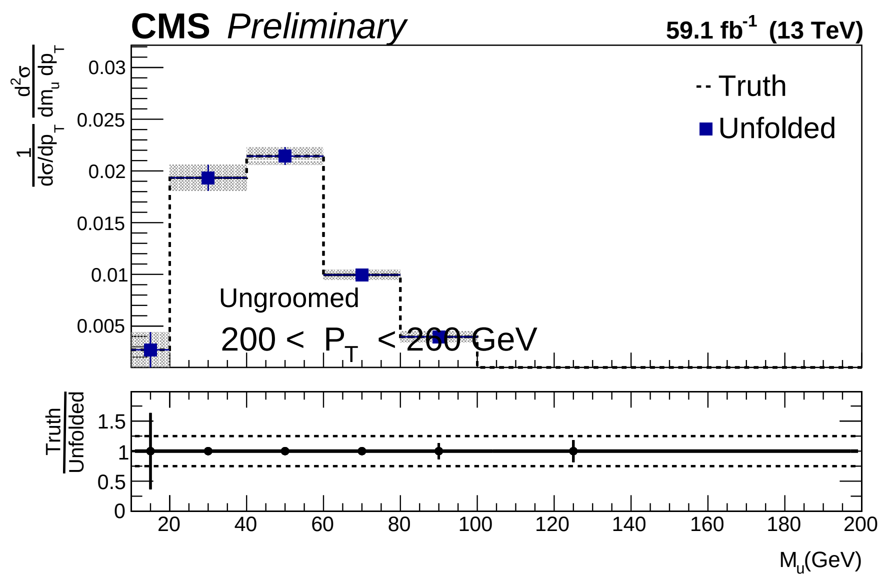

Figure 226: Closure test of ungroomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

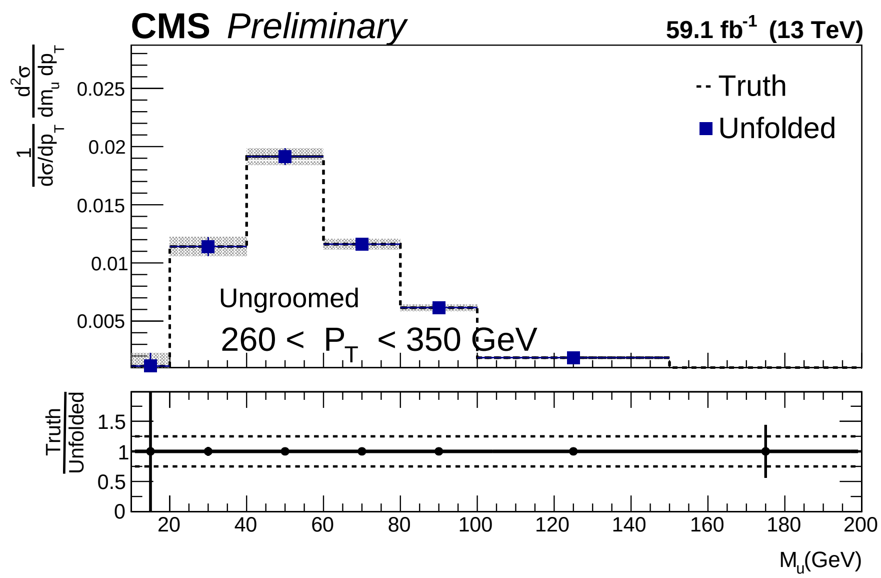

Figure 227: Closure test of ungroomed reconstructed Monte Carlo, p_T 260-350 GeV.

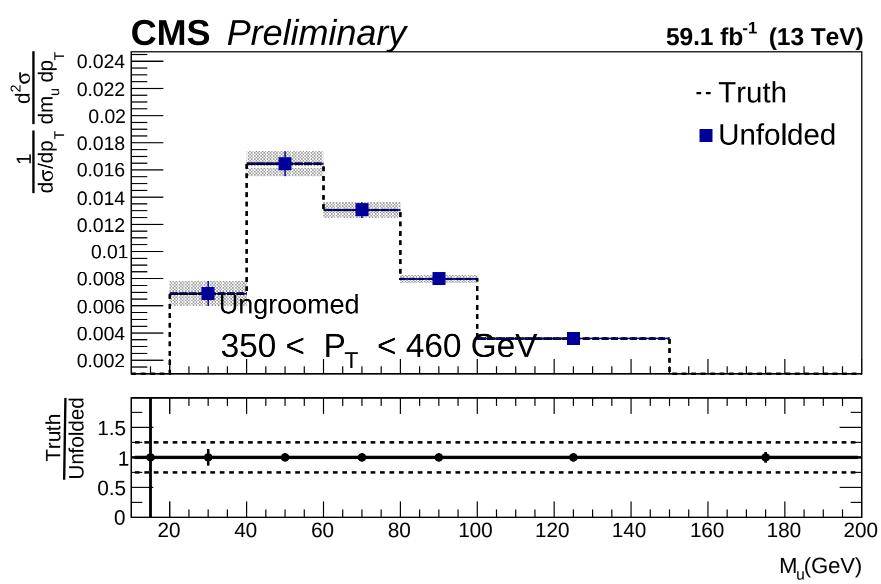

Figure 228: Closure test of ungroomed reconstructed Monte Carlo, p_T 350-460 GeV.

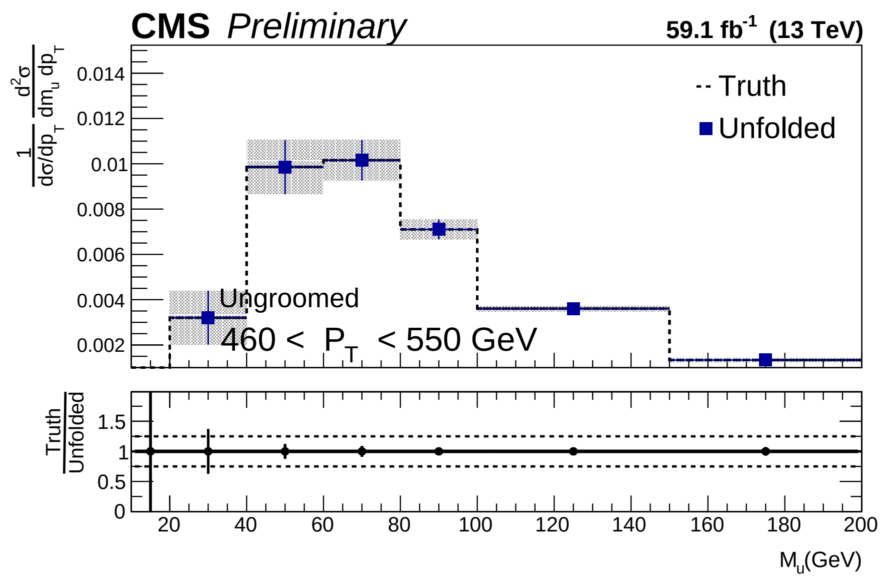

Figure 229: Closure test of ungroomed reconstructed Monte Carlo, p_T 460-550 GeV.

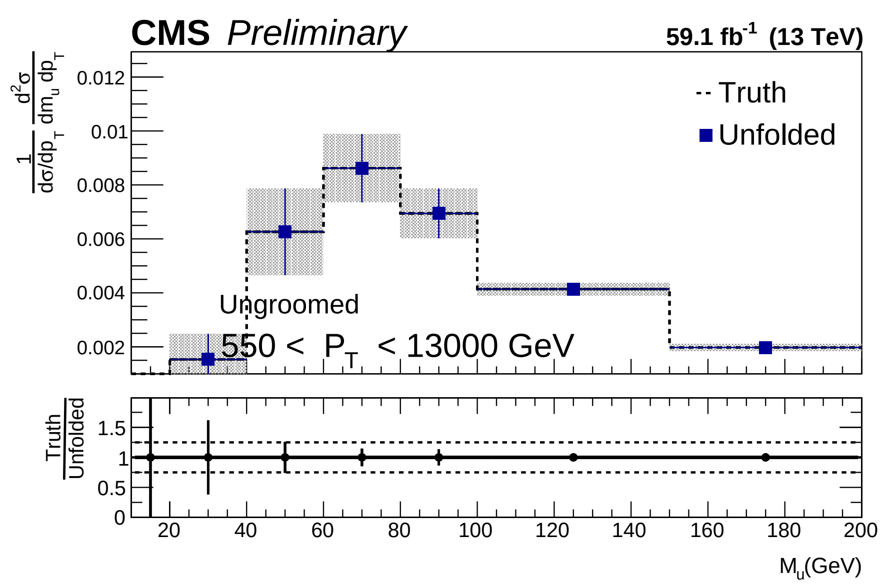

Figure 230: Closure test of ungroomed reconstructed Monte Carlo, p_T 550-13000 GeV.

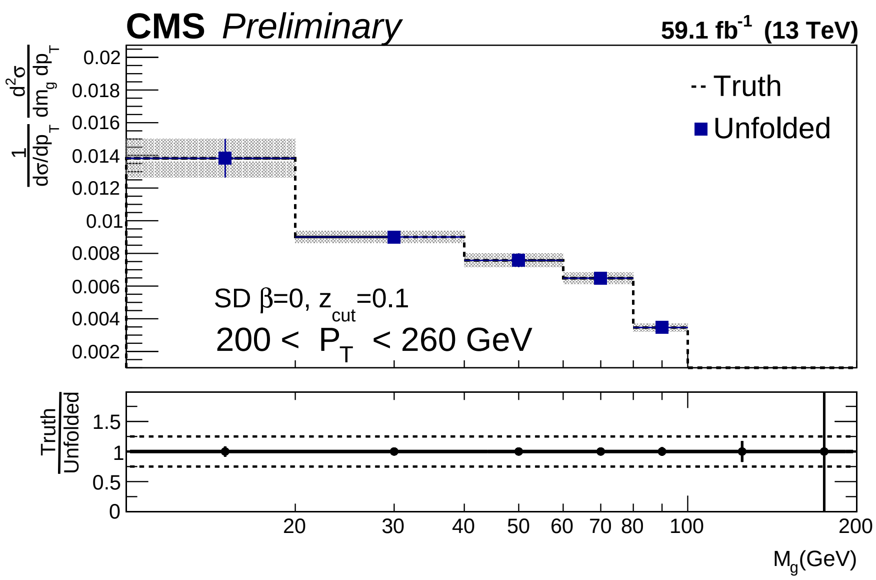

Figure 231: Closure test of groomed reconstructed Monte Carlo, p_{T} 200-260 GeV.

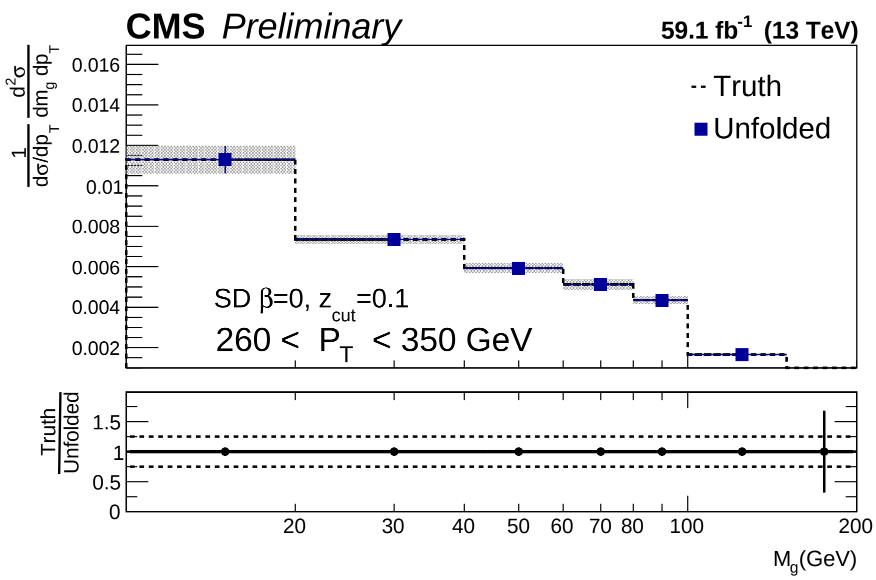

Figure 232: Closure test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

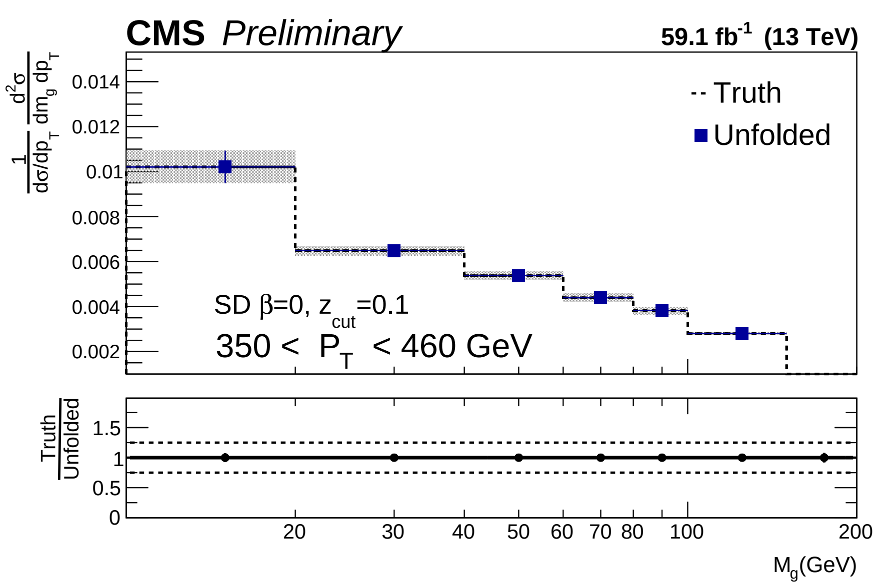

Figure 233: Closure test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

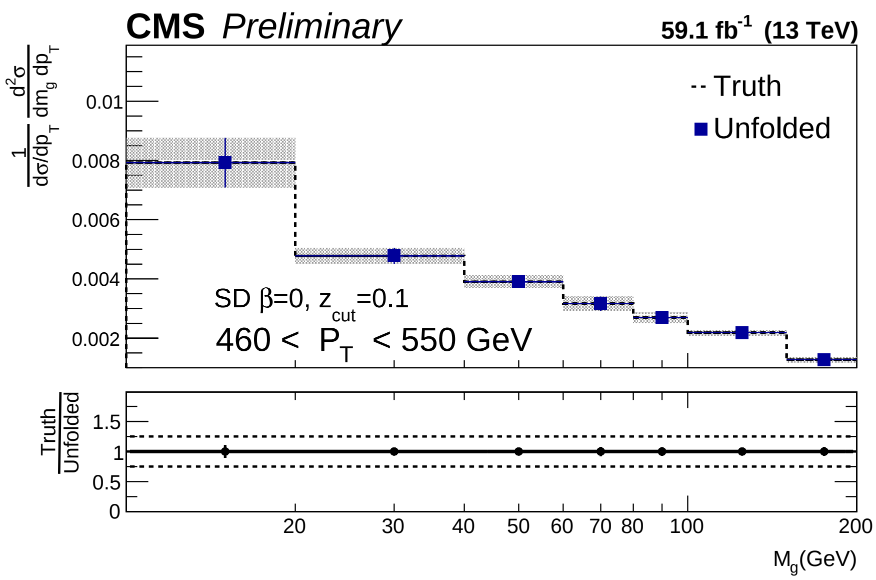

Figure 234: Closure test of groomed reconstructed Monte Carlo, p_{T} 460-550 GeV.

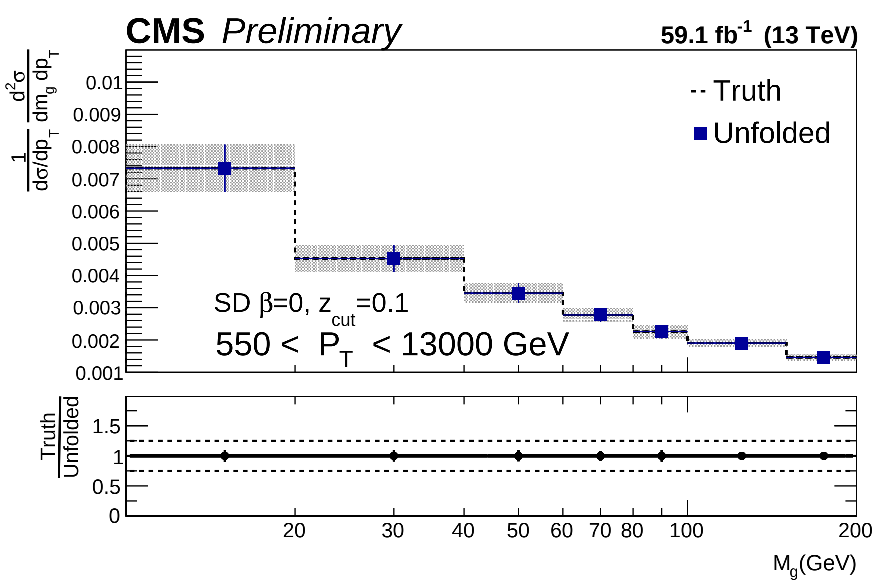

Figure 235: Closure test of groomed reconstructed Monte Carlo, p_{T} 550-13000 GeV.

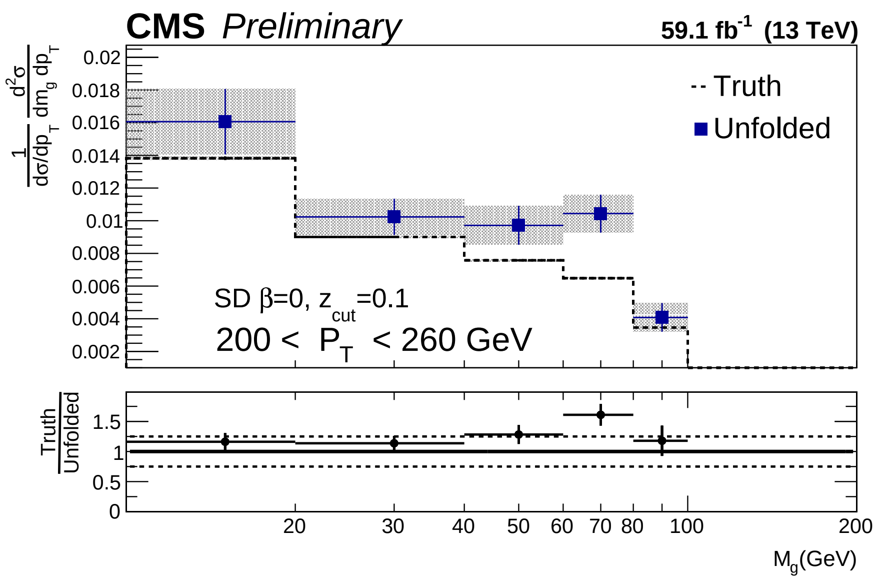

Figure 236: Physical bias test of groomed reconstructed Monte Carlo, p_T 200-260 GeV.

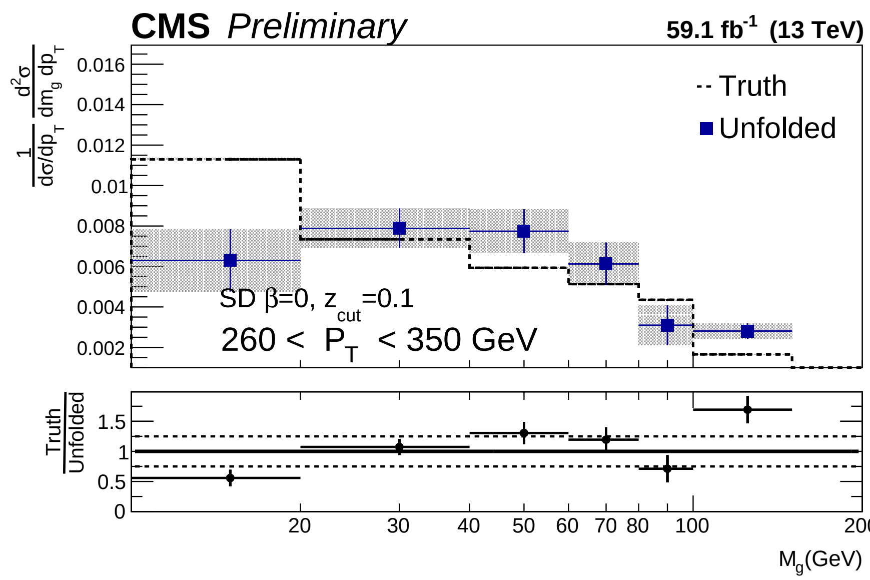

Figure 237: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 260-350 GeV.

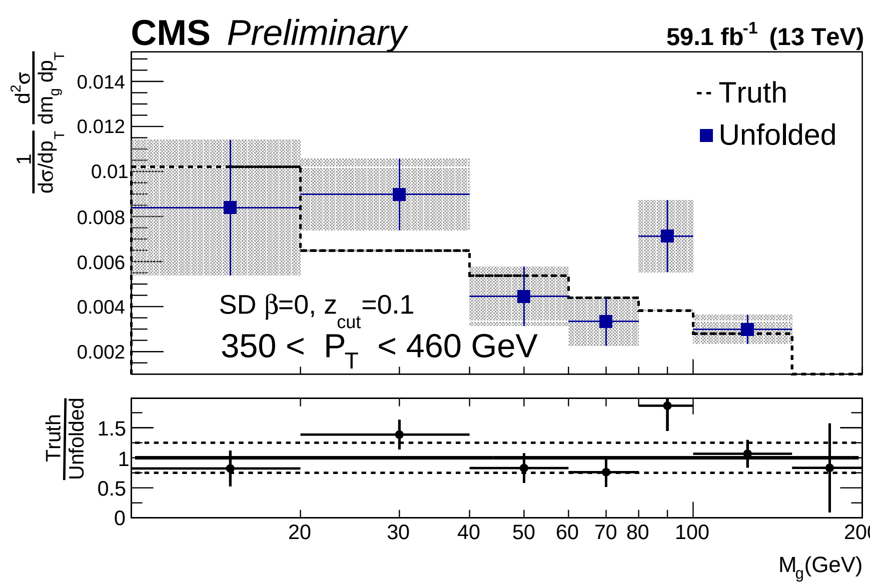

Figure 238: Physical bias test of groomed reconstructed Monte Carlo, p_{T} 350-460 GeV.

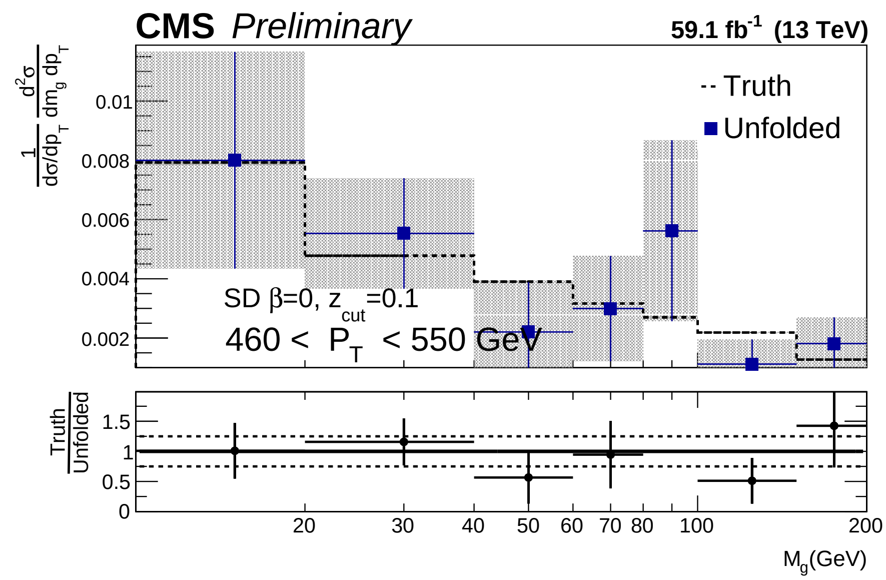

Figure 239: Physical bias test of groomed reconstructed Monte Carlo, p_T 460-550 GeV.

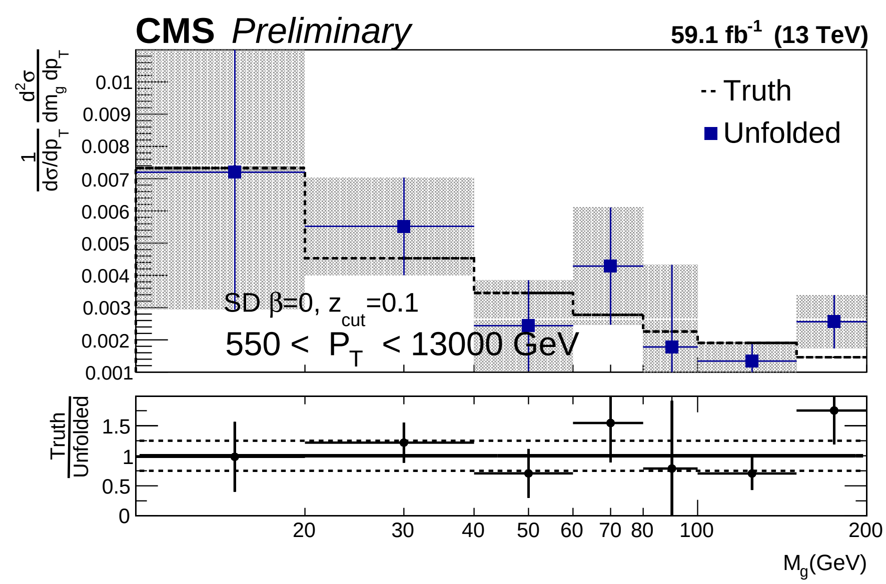

Figure 240: Physical bias test of groomed reconstructed Monte Carlo, p_T 550-13000 GeV.

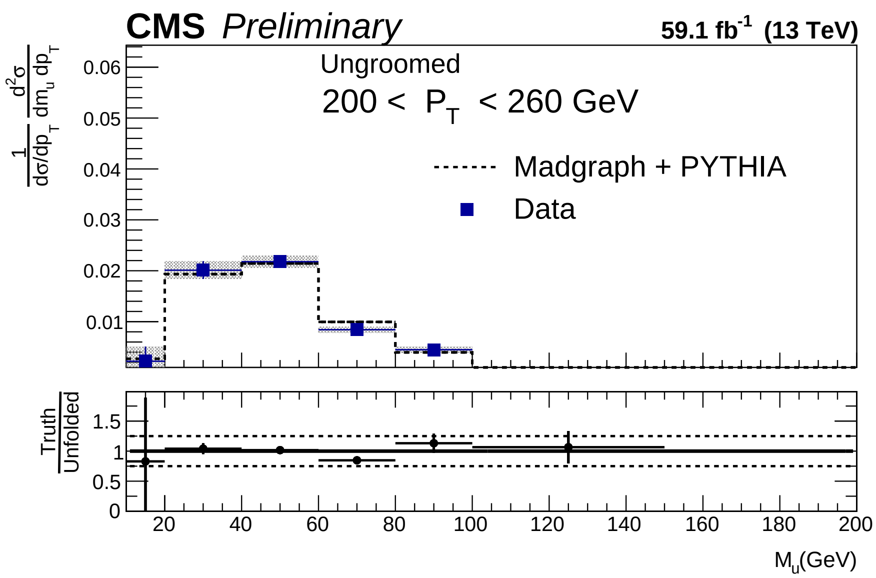

Figure 241: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 200-260 GeV.

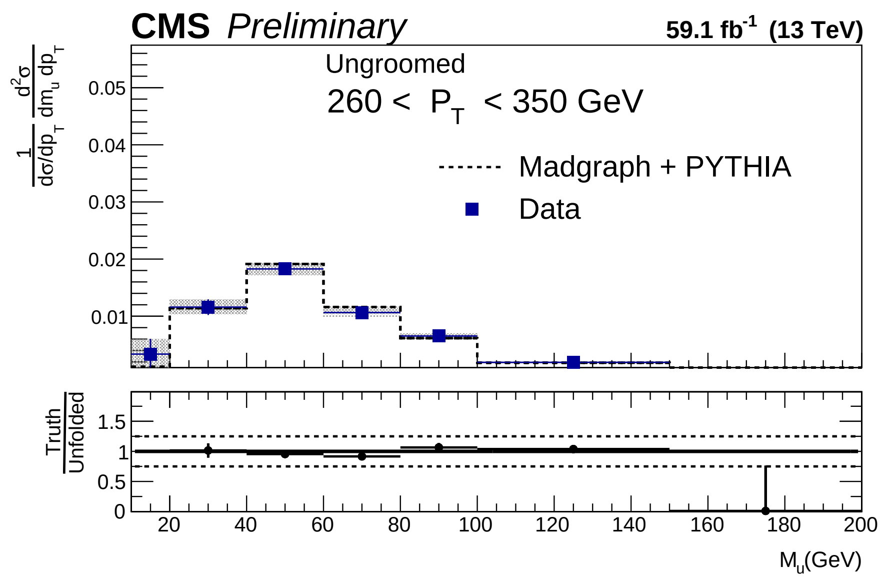

Figure 242: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 260-350 GeV.

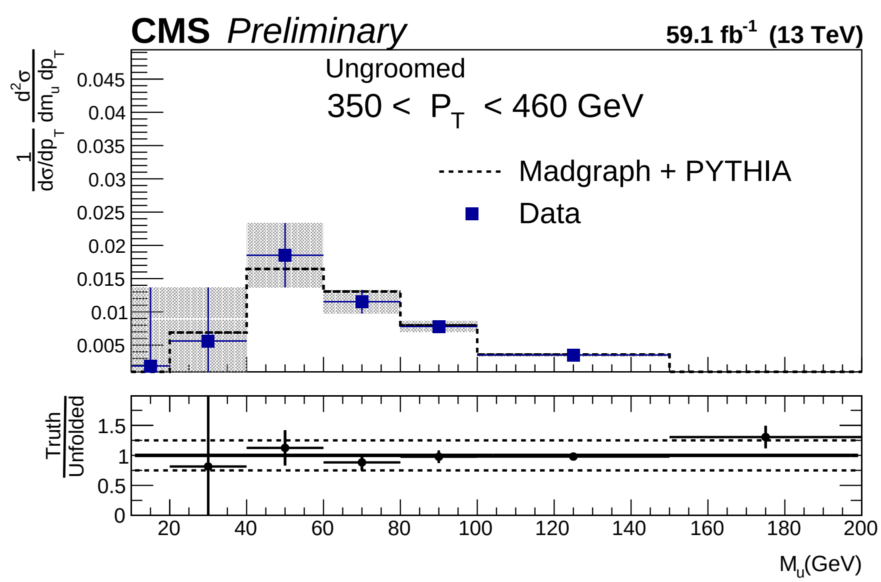

Figure 243: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 350-460 GeV.

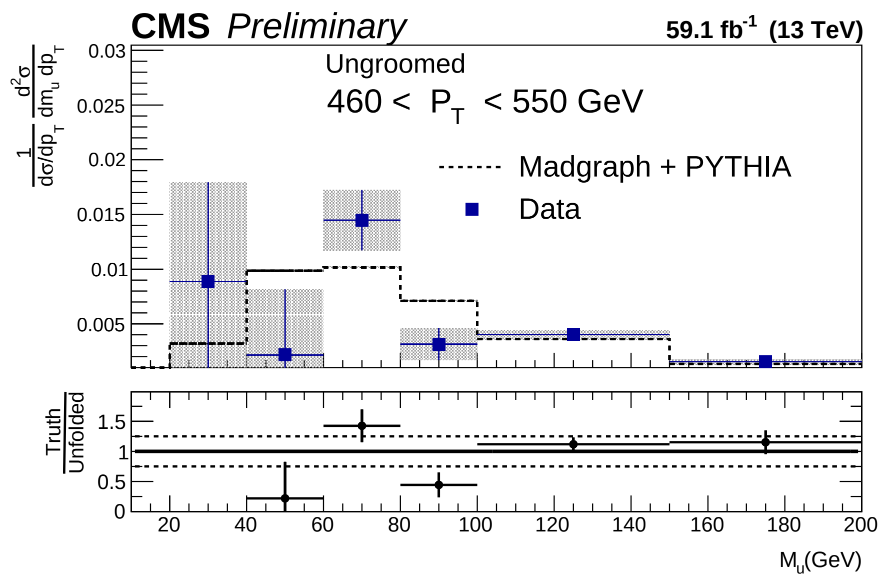

Figure 244: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 460-550 GeV.

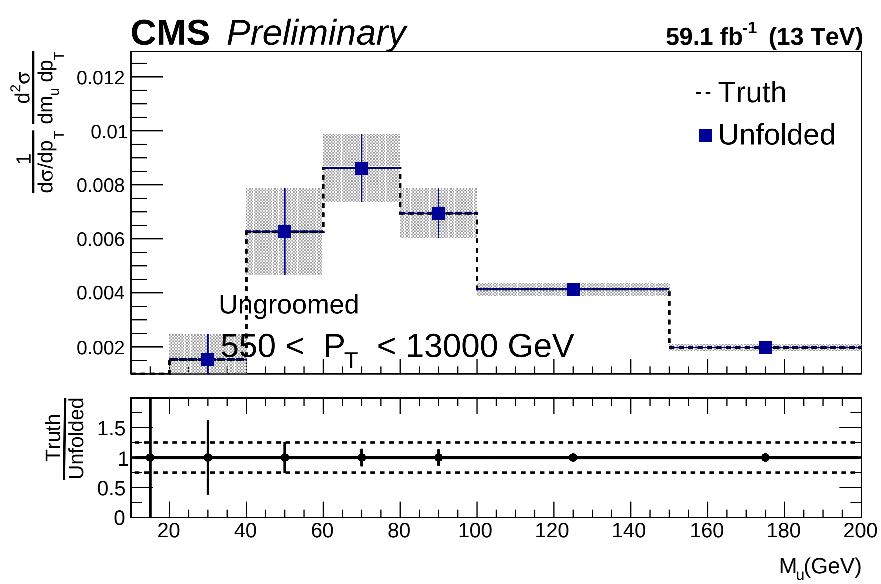

Figure 245: Normalized cross section results with respect to jet mass for ungroomed jets, p_{T} 550-13000 GeV.

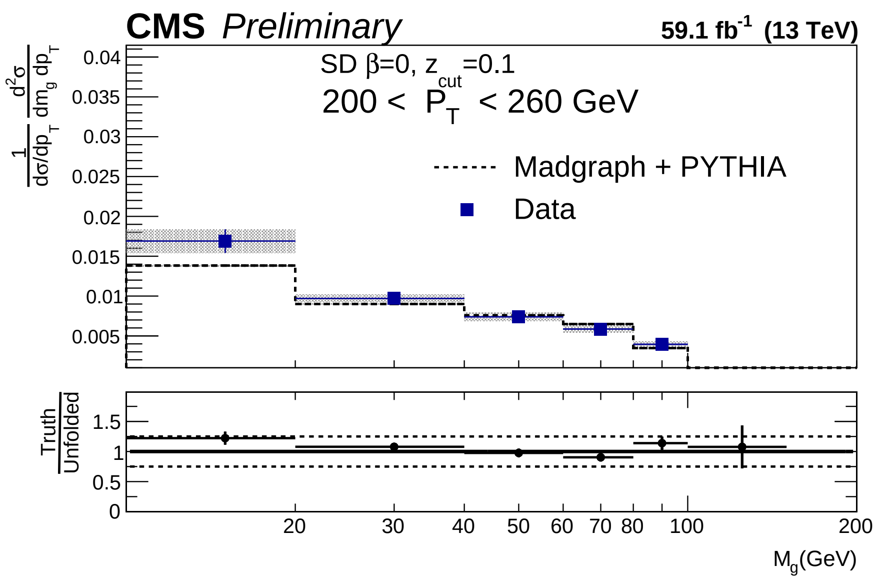

Figure 246: Normalized cross section results with respect to jet mass for groomed jets, p_{T} 200-260 GeV.

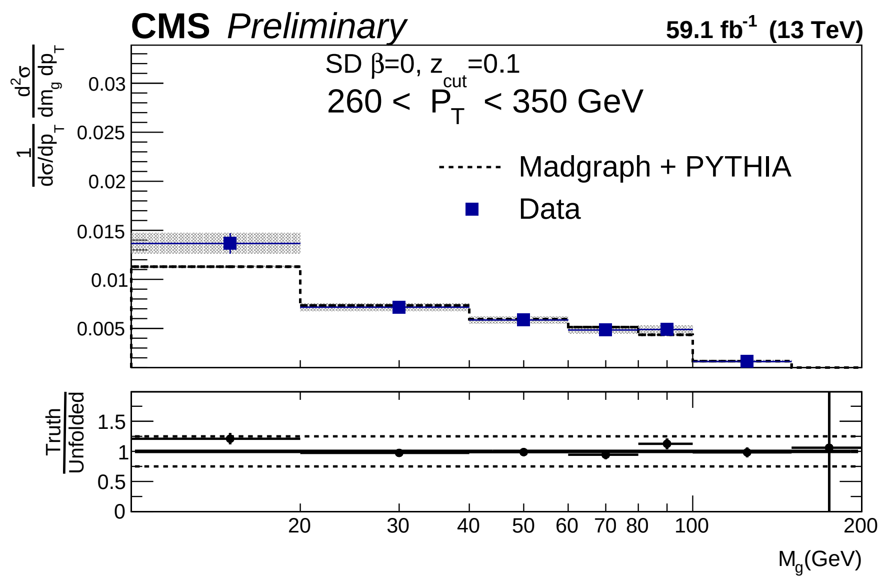

Figure 247: Normalized cross section results with respect to jet mass for groomed jets, p_T 260-350 GeV.

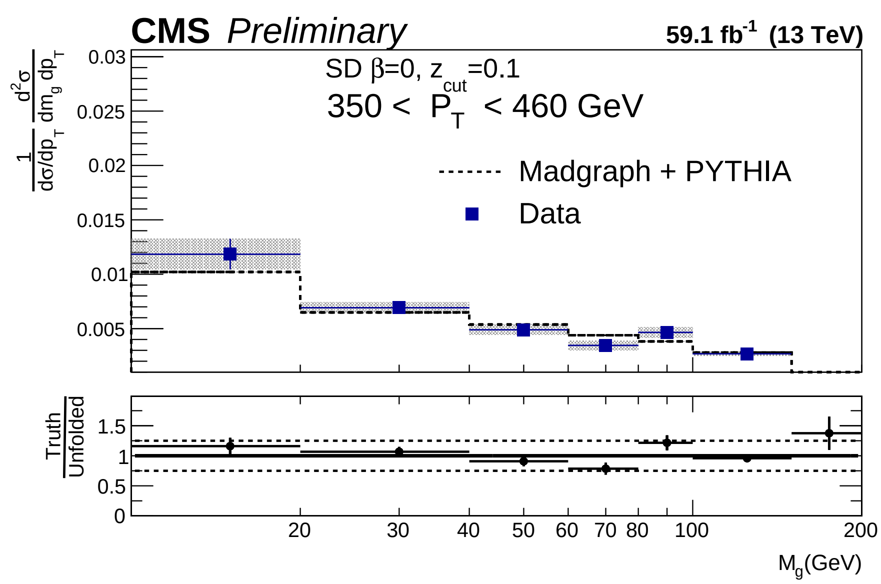

Figure 248: Normalized cross section results with respect to jet mass for groomed jets, p_T 350-460 GeV.

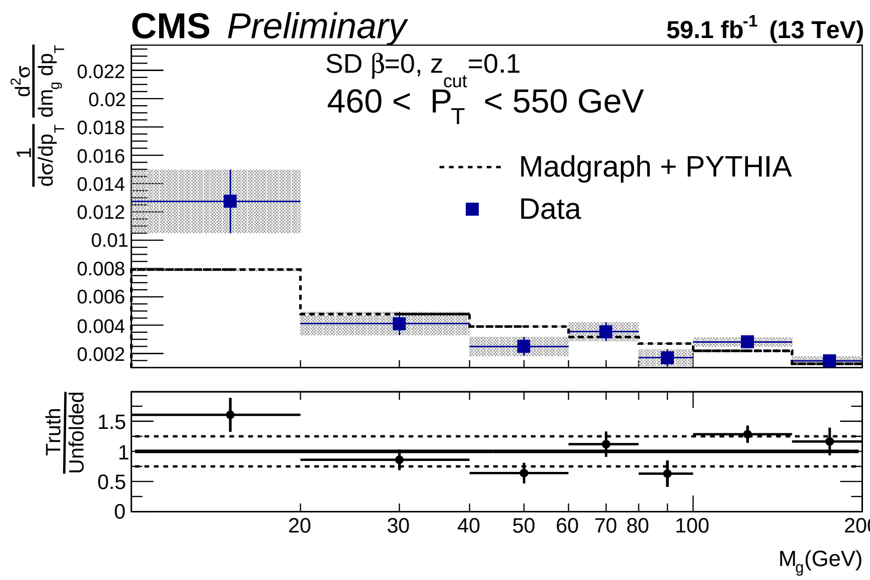

Figure 249: Normalized cross section results with respect to jet mass for groomed jets, p_T 460-550 GeV.

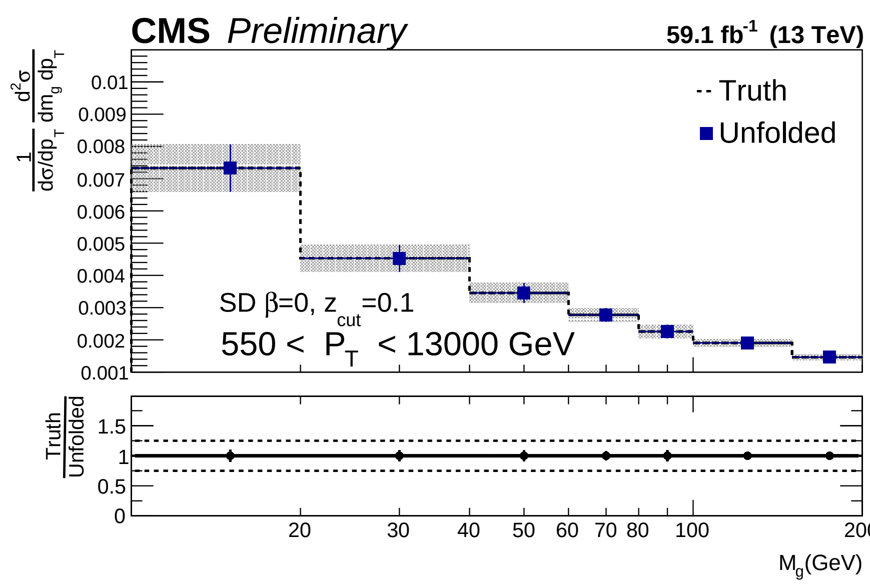

Figure 250: Normalized cross section results with respect to jet mass for groomed jets, p_T 550-13000 GeV.

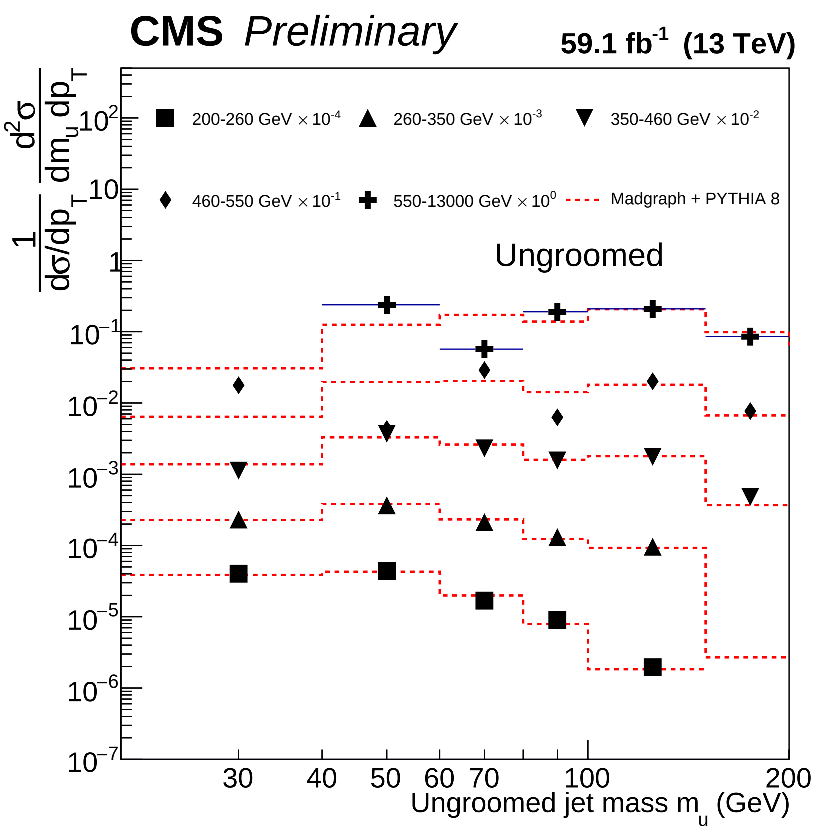

Figure 251: Results for ungroomed reconstructed unfolding with jet mass.

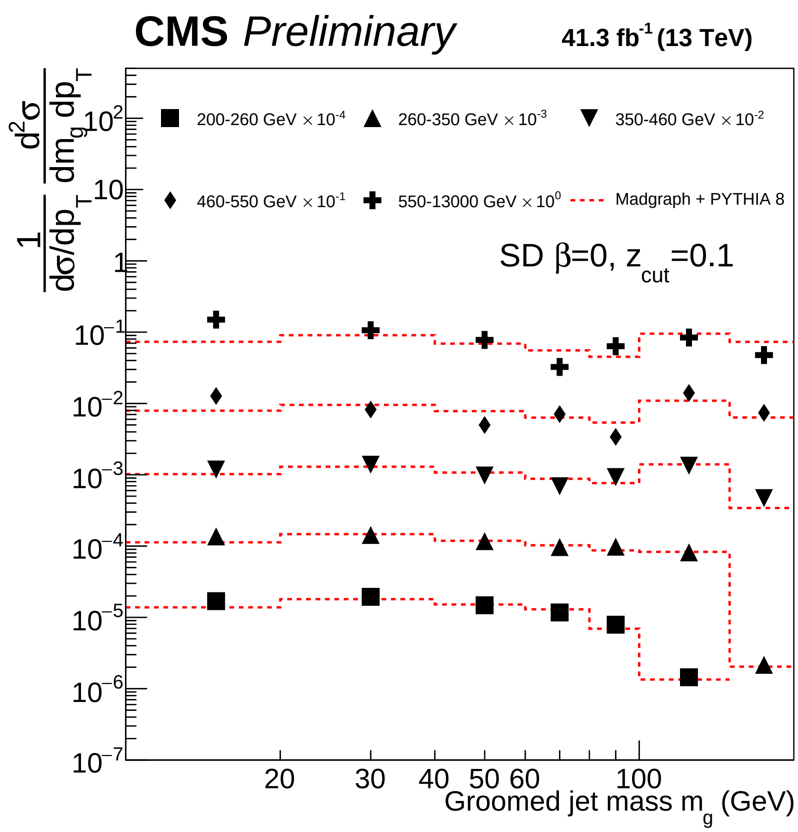

Figure 252: Results for groomed reconstructed unfolding with jet mass.

The resulting correlation matrix without systematic uncertainties for the ungroomed jets is shown in Fig. 253, and for the groomed jets is shown in Fig. 254. The same figures with systematic uncertainties are shown for the ungroomed jets in Fig. 255, and for the groomed jets is shown in Fig. 256.

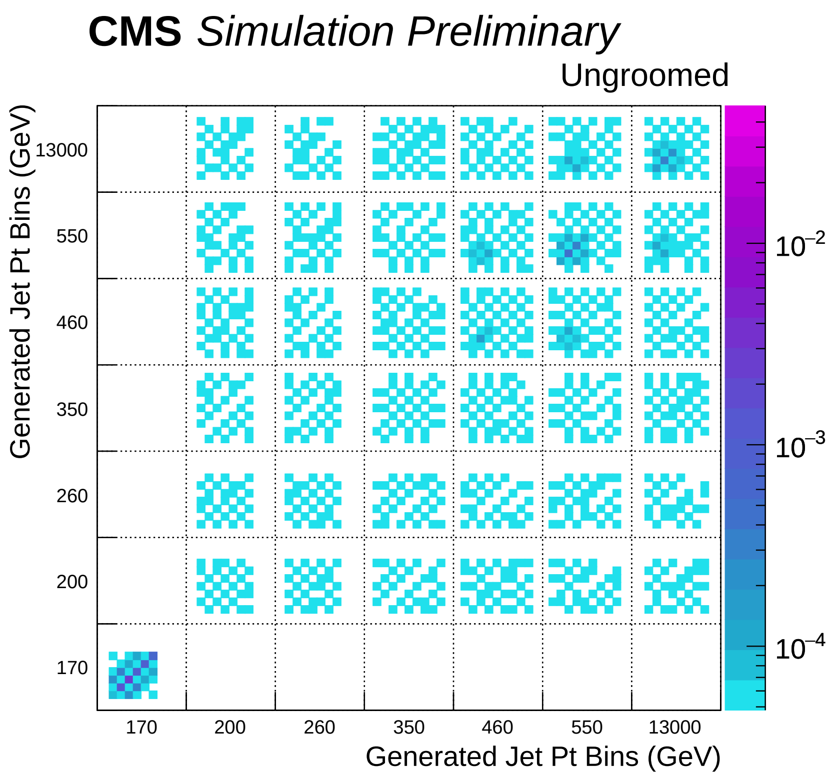

Figure 253: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,10,20,. . . 200 GeV), while the p_T is indexed by the major blocks (200,350,. . . 13000 GeV).

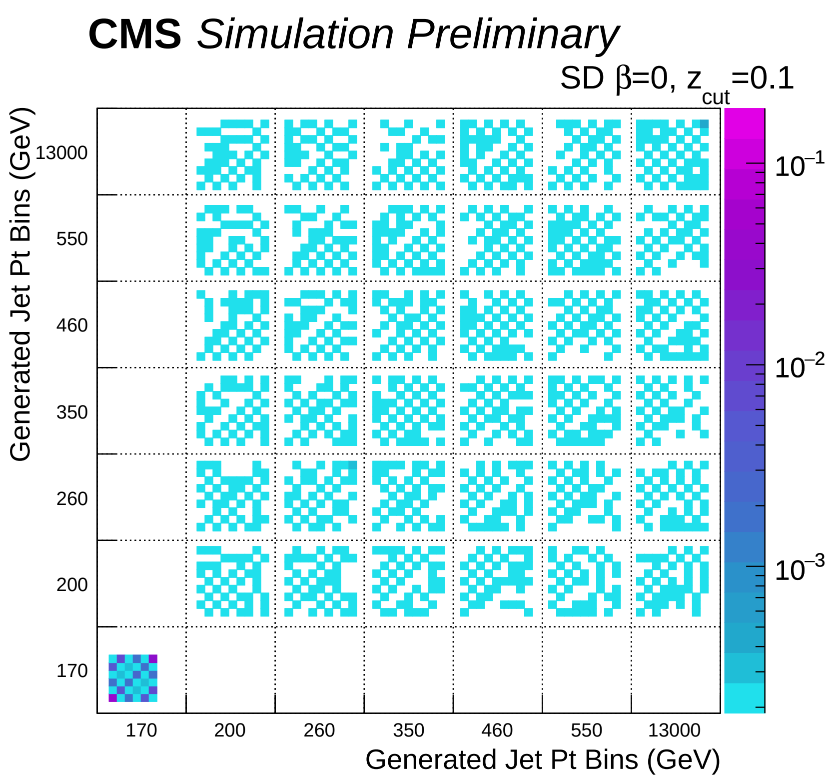

Figure 254: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 760 GeV).

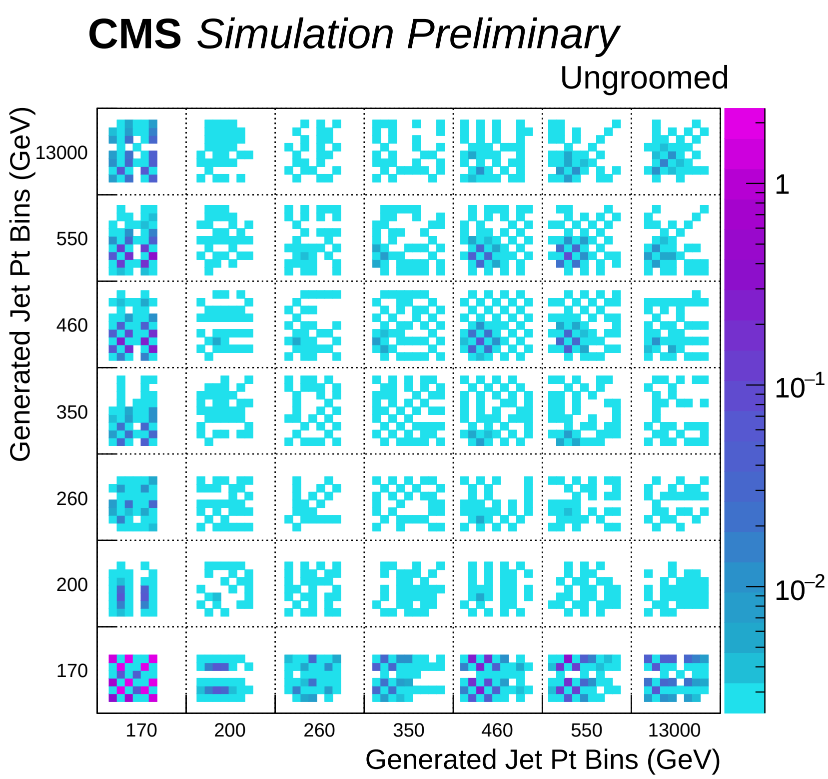

Figure 255: Correlation matrix for ungroomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,... 200 GeV), while the p_T is indexed by the major blocks (200,260,... 13000 GeV).

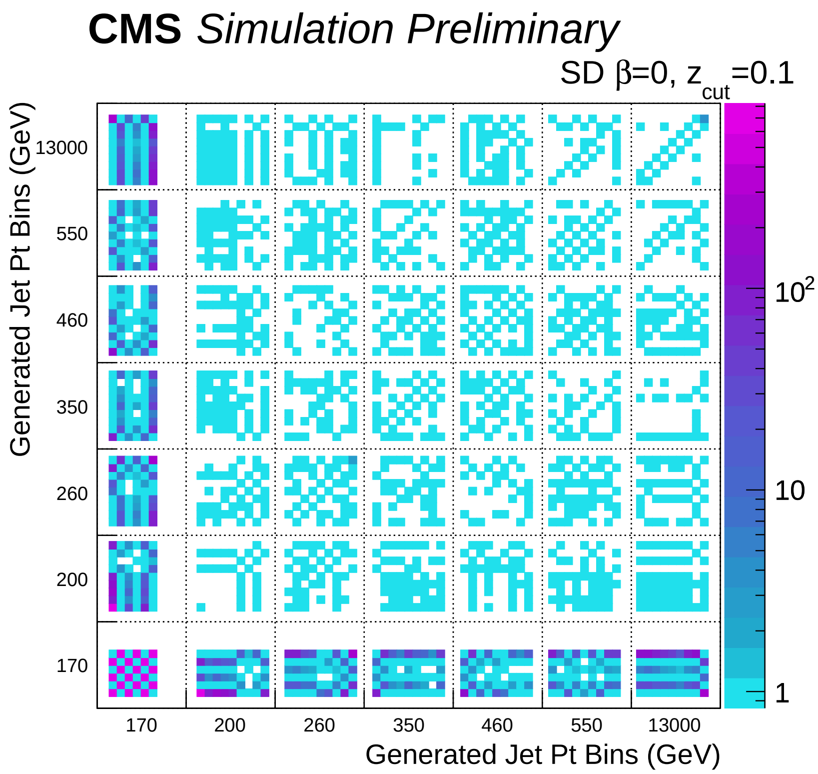

Figure 256: Correlation matrix for groomed jets. The bins are set such that the jet mass is indexed by the minor blocks (0,20,40,. . . 200 GeV), while the p_T is indexed by the major blocks (200,340,. . . 0 GeV).

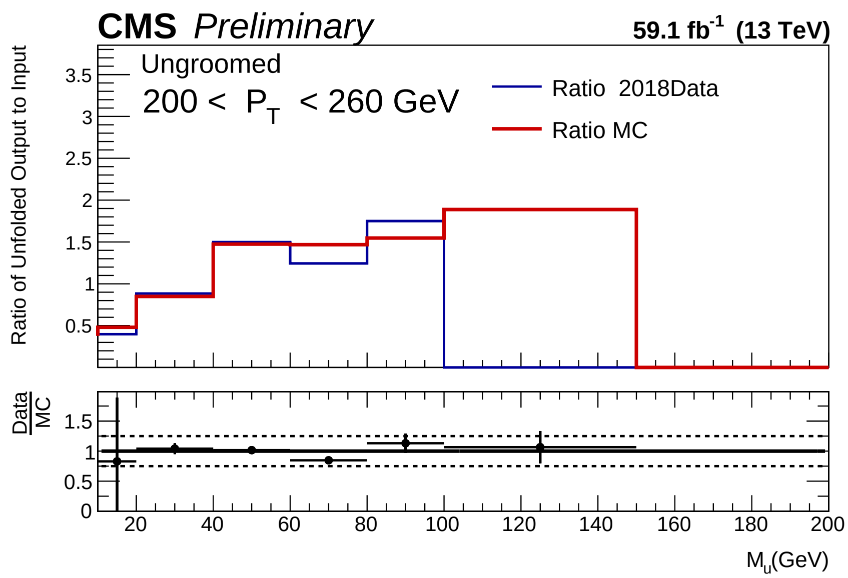

Figure 257: Ratio of unfolded over raw data and MC for ungroomed jets, p_{T} 200-260 GeV.

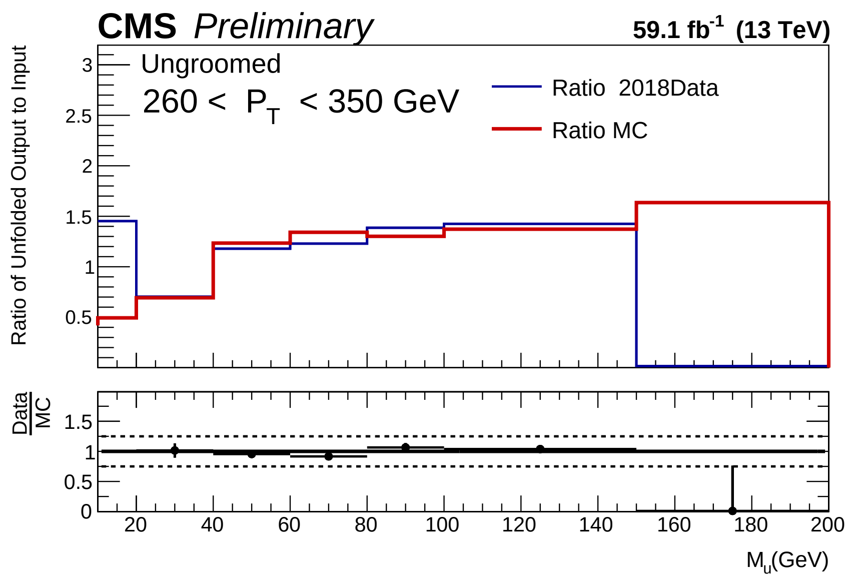

Figure 258: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 260-350 GeV.

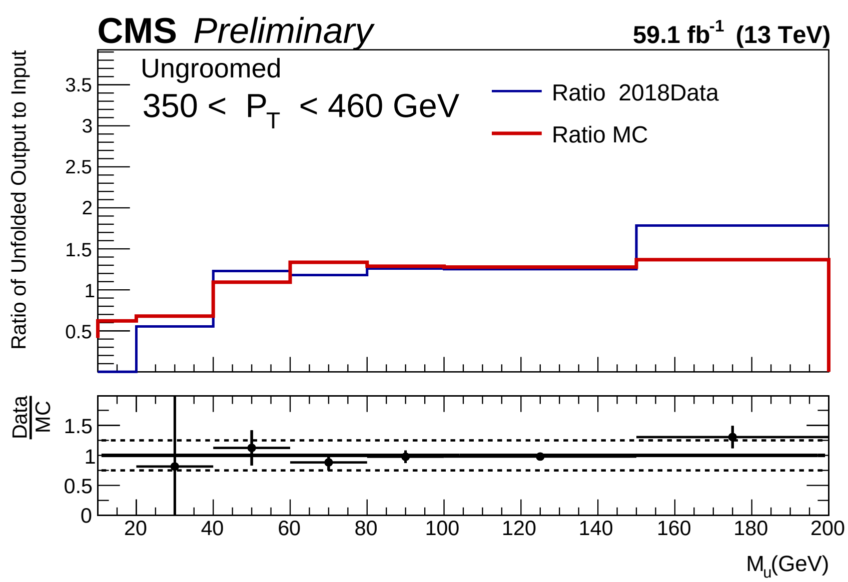

Figure 259: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 350-460 GeV.

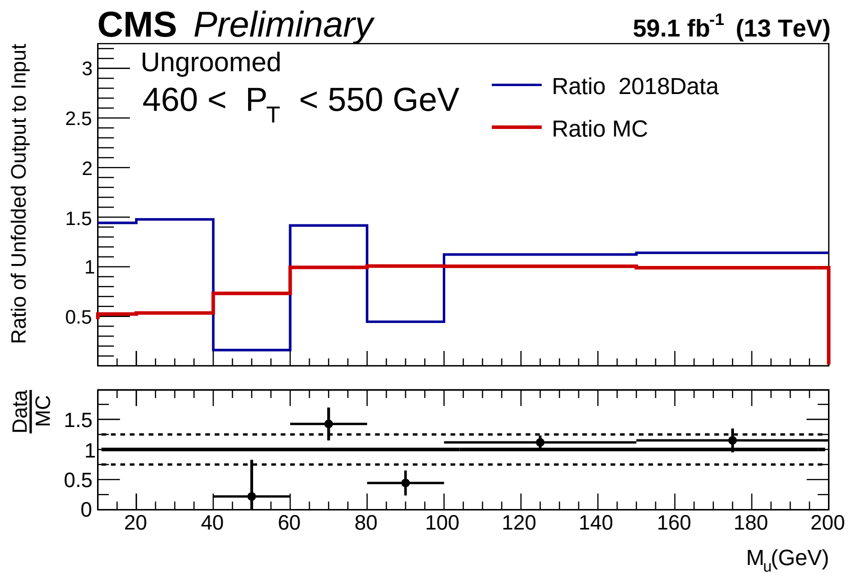

Figure 260: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 460-550 GeV.

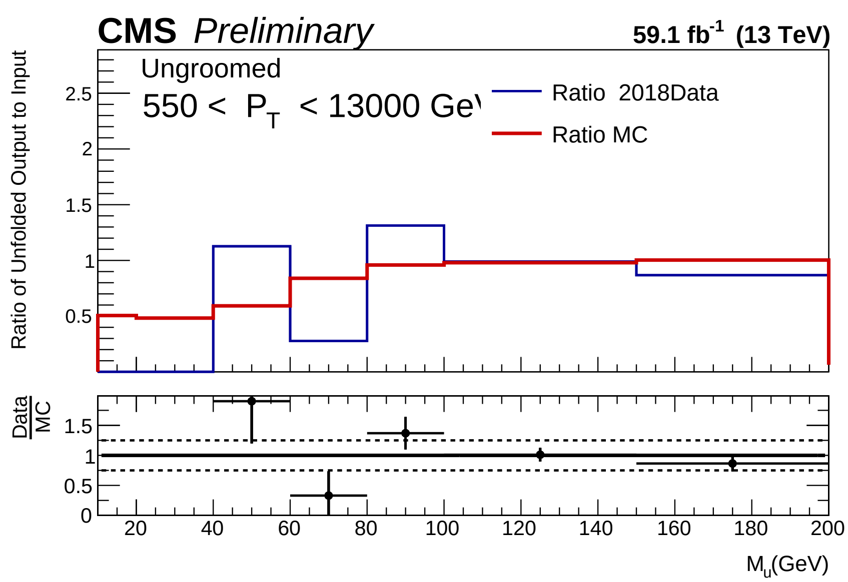

Figure 261: Ratio of unfolded over raw data and MC for ungroomed jets, p_T 550-13000 GeV.

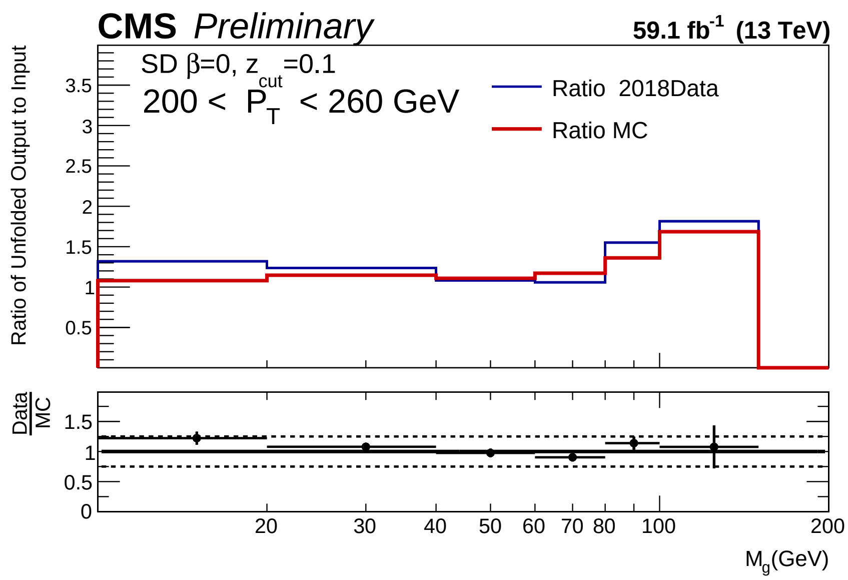

Figure 262: Ratio of unfolded over raw data and MC for groomed jets, p_{T} 200-260 GeV.

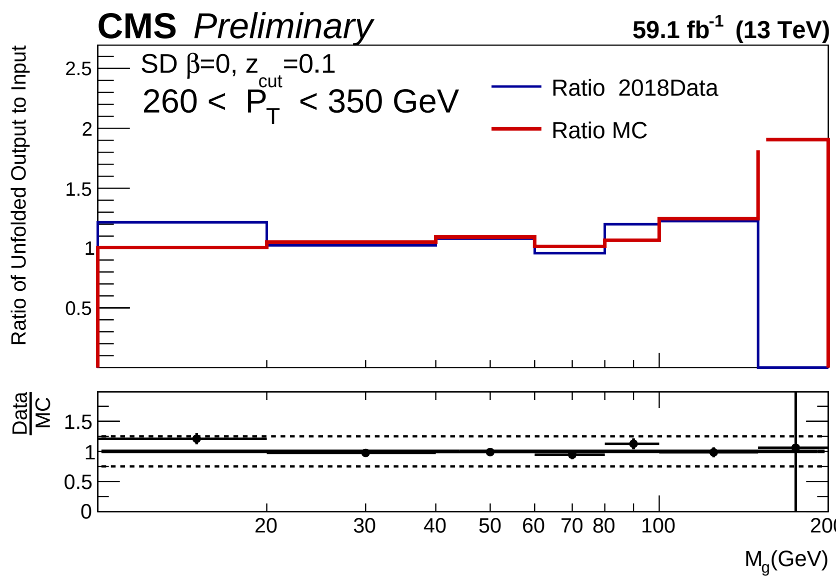

Figure 263: Ratio of unfolded over raw data and MC for groomed jets, p_T 260-350 GeV.

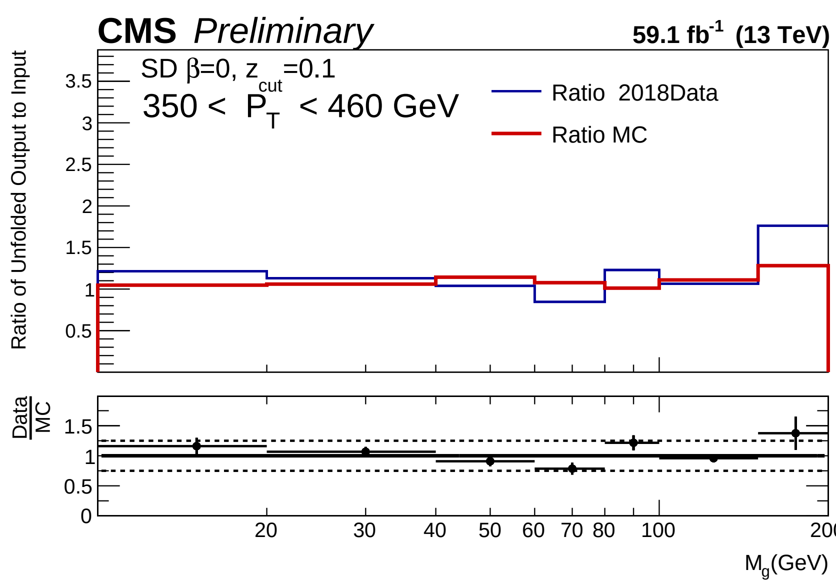

Figure 264: Ratio of unfolded over raw data and MC for groomed jets, p_T 350-460 GeV.

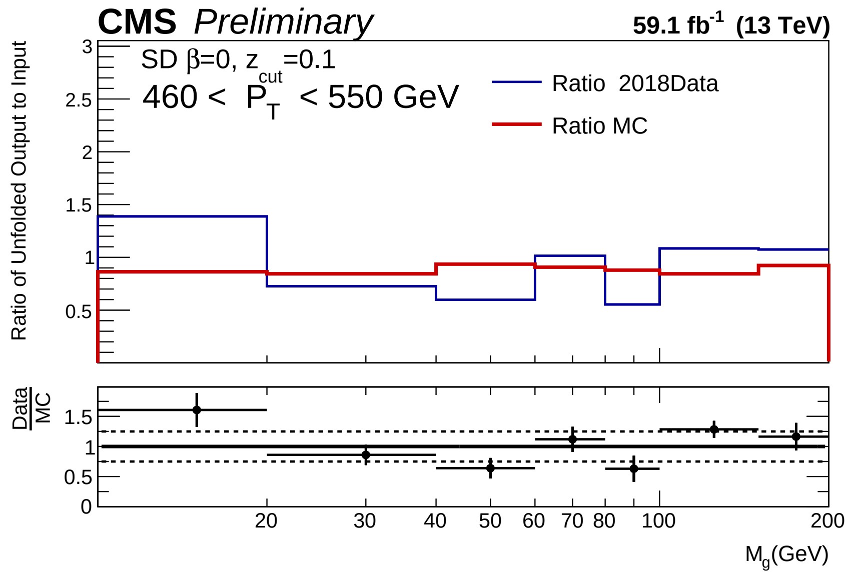

Figure 265: Ratio of unfolded over raw data and MC for groomed jets, p_T 460-550 GeV.

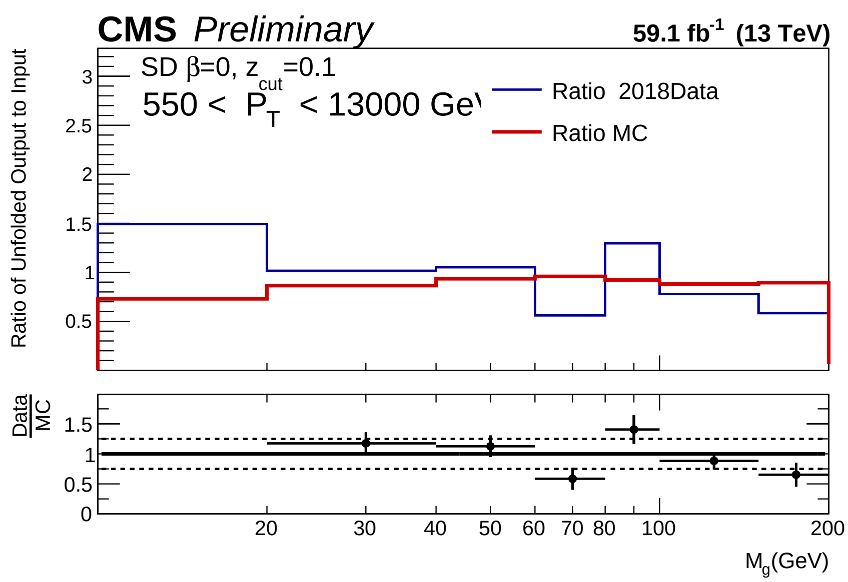

Figure 266: Ratio of unfolded over raw data and MC for groomed jets, p_T 550-13000 GeV.

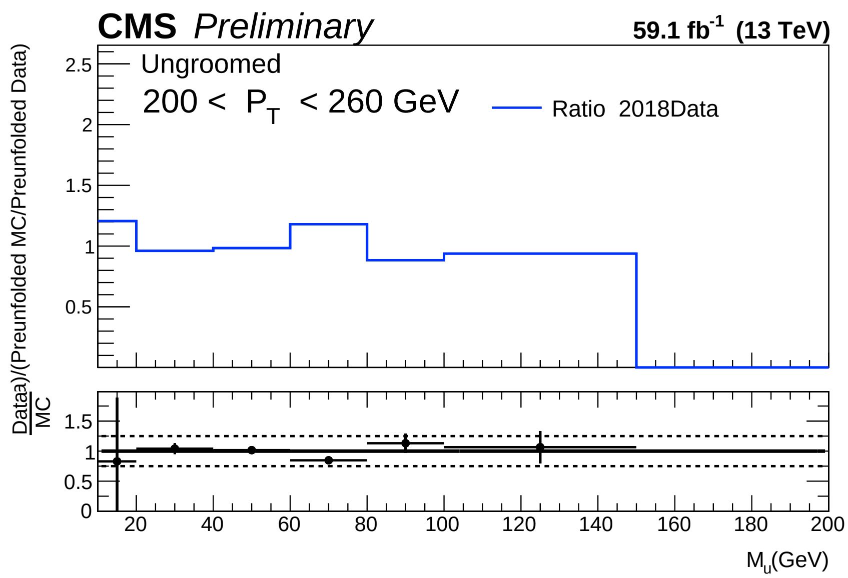

Figure 267: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_T 200-260 GeV.

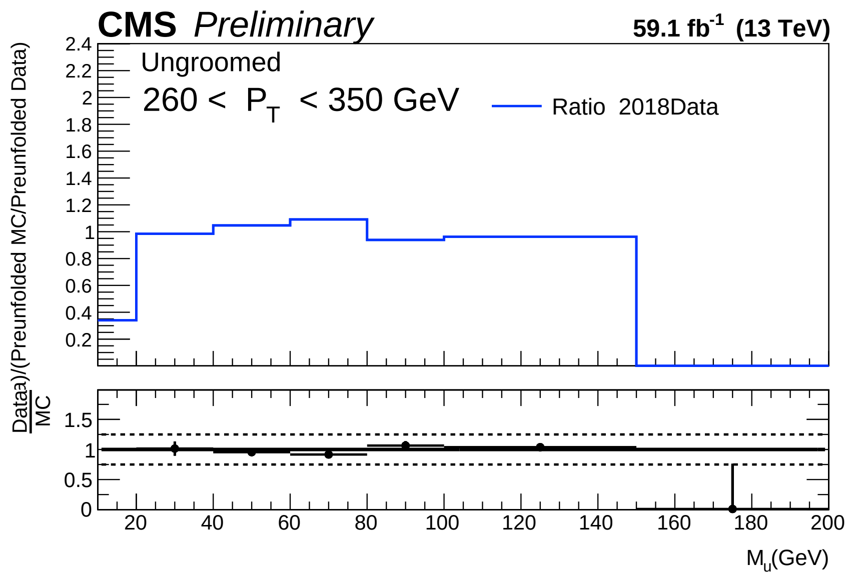

Figure 268: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 260-350 GeV.

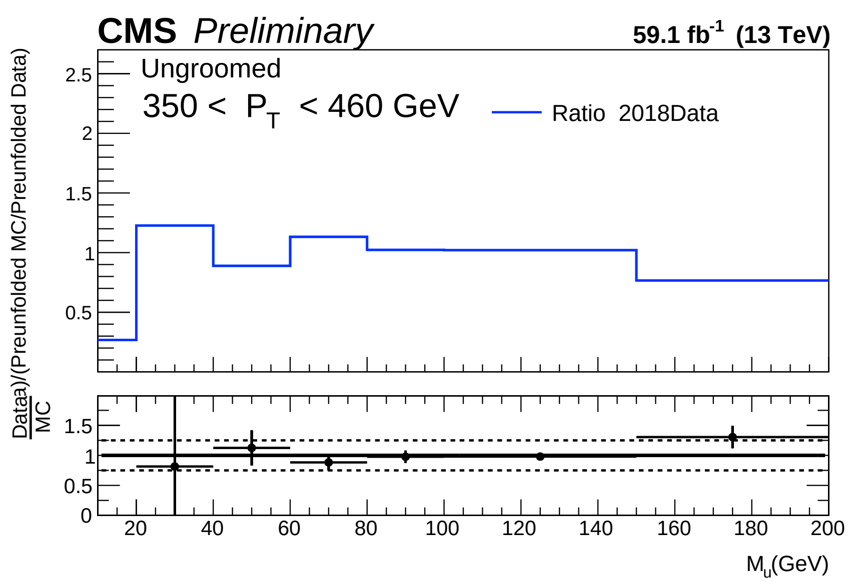

Figure 269: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 350-460 GeV.

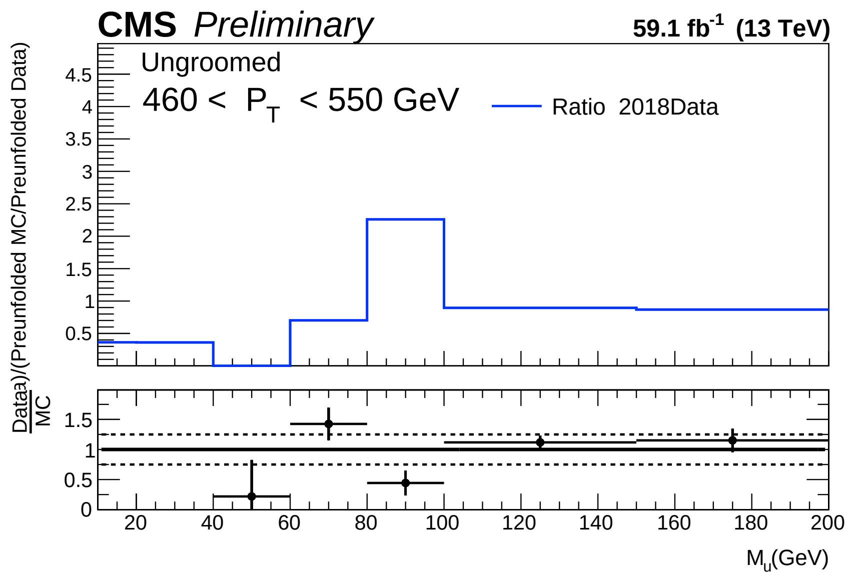

Figure 270: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 460-550 GeV.

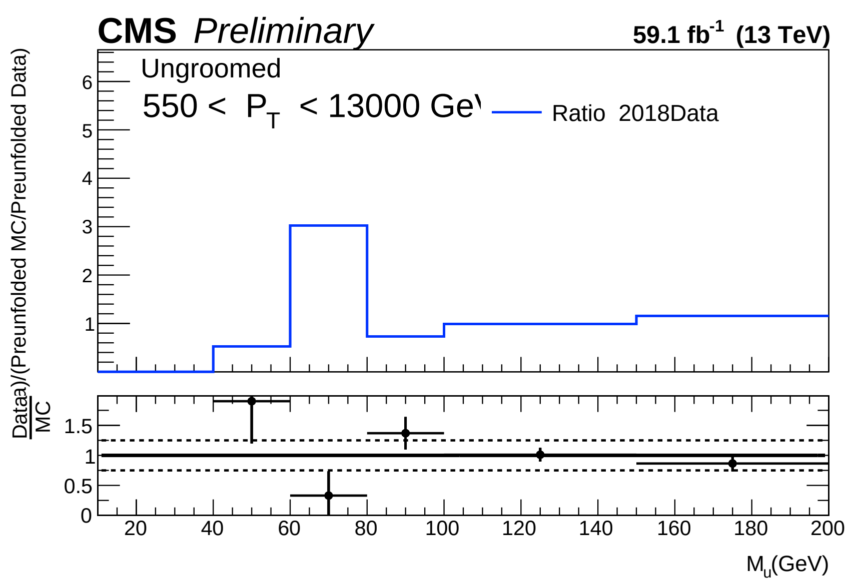

Figure 271: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for ungroomed jets, p_{T} 550-13000 GeV.

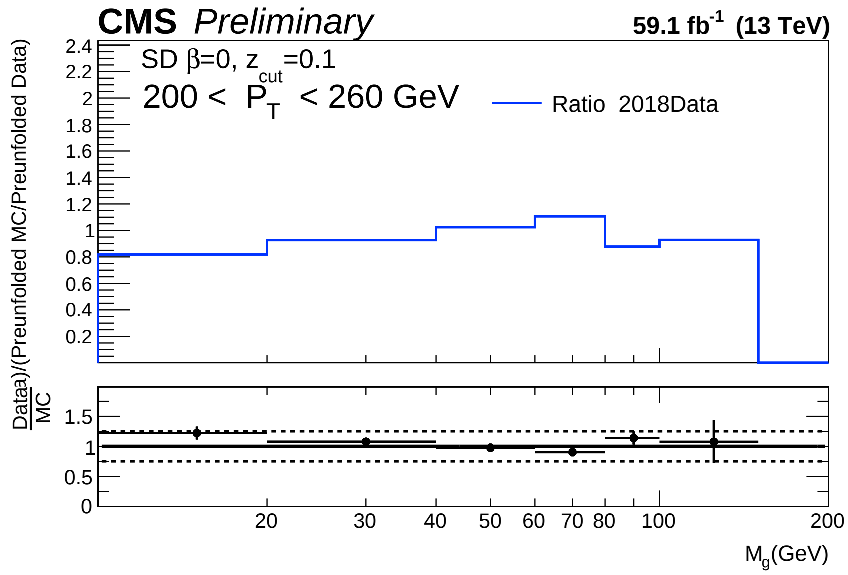

Figure 272: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 200-260 GeV.

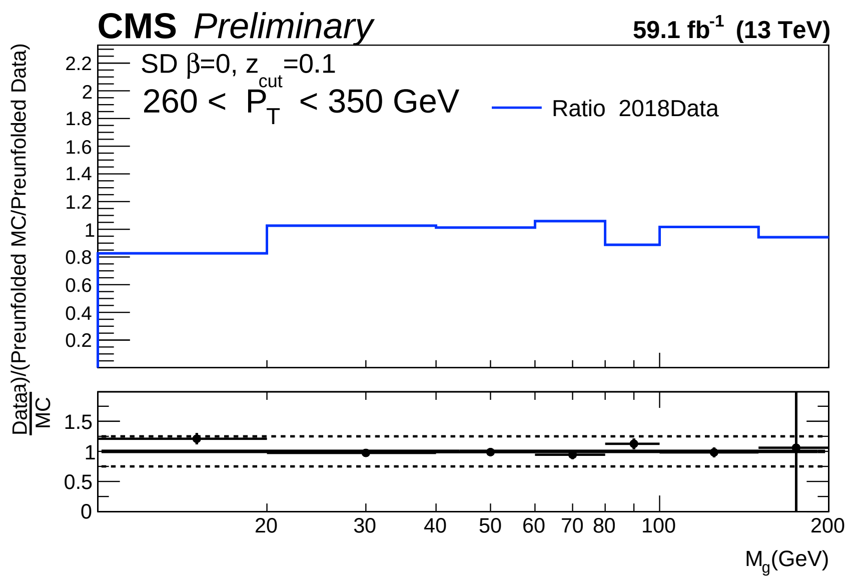

Figure 273: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 260-350 GeV.

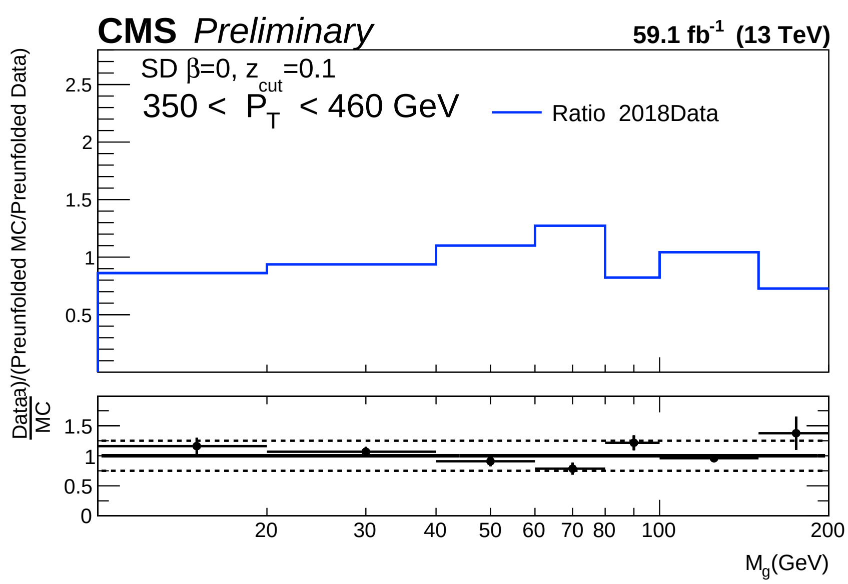

Figure 274: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 350-460 GeV.

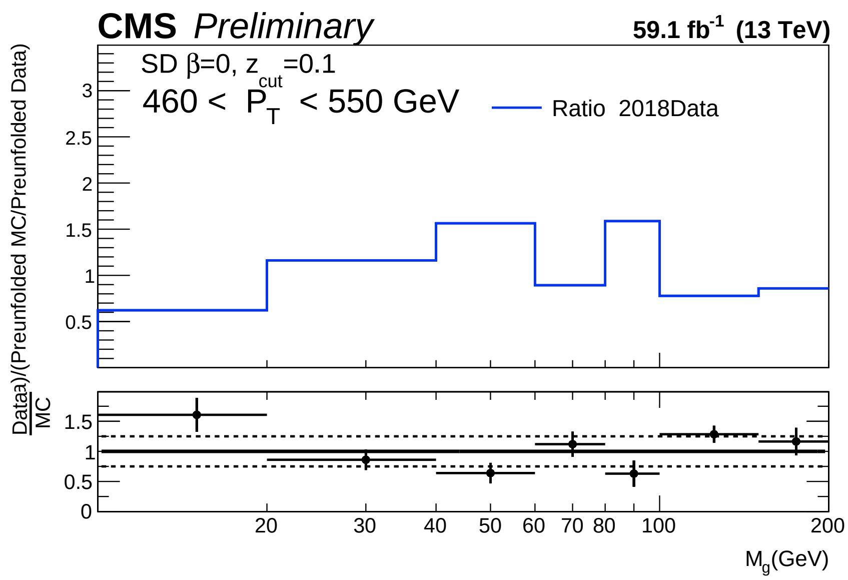

Figure 275: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 460-550 GeV.

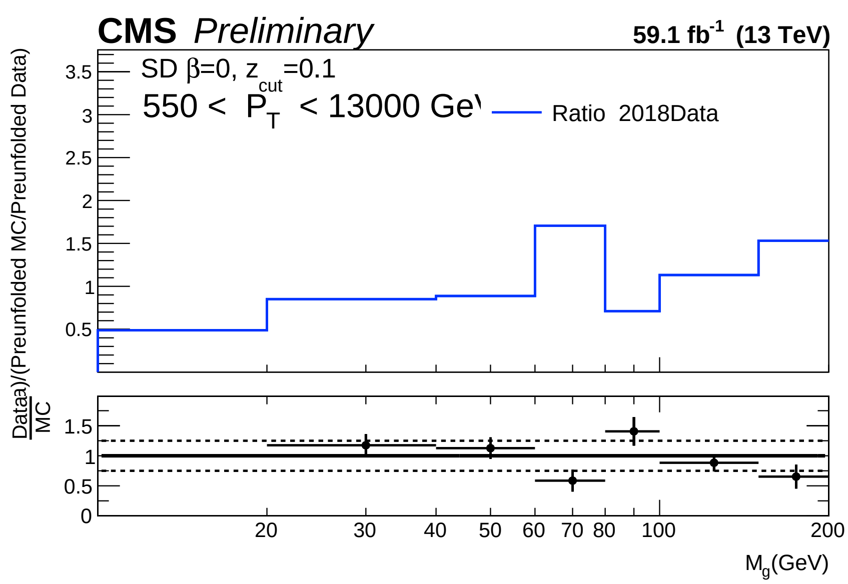

Figure 276: Ratio of generator truth over unfolded data, divided by the ratio of raw MC over raw data for groomed jets, p_{T} 550-13000 GeV.

Appendix A

Conclusion

A.1 Conclusion

Theoretical calculations are being prepared by our collegues and will soon be compared to the soft dropped jet unfolding results presented in this thesis.

Bibliography

[1] Vardan Khachatryan et al. Measurement of the inclusive 3-jet production differential cross section in proton–proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range. *Eur. Phys. J.*, C75(5):186, 2015.

[2] modellinginvisible.org standard model description. `https://www.modellinginvisible.org/standard-model/`. Accessed: 2019-08-06.

[3] Frédéric A. Dreyer, Gavin P. Salam, and Grégory Soyez. The Lund Jet Plane. *JHEP*, 12:064, 2018.

[4] Thomas Becher, Alessandro Broggio, and Andrea Ferroglia. Introduction to Soft-Collinear Effective Theory. *Lect. Notes Phys.*, 896:pp.1–206, 2015.

[5] Serguei Chatrchyan et al. Studies of Jet Mass in Dijet and W/Z + Jet Events. *JHEP*, 05:090, 2013.

[6] Albert M. Sirunyan et al. Measurements of the differential jet cross section as a function of the jet mass in dijet events from proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 11:113, 2018.

[7] Morad Aaboud et al. Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 121(9):092001, 2018.

[8] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008.

[9] Y. Ashie et al. A Measurement of atmospheric neutrino oscillation parameters by SUPER-KAMIOKANDE I. *Phys. Rev.*, D71:112005, 2005.

[10] Andrew J. Larkoski. An Unorthodox Introduction to QCD. 2017.

[11] Lily Asquith et al. Jet Substructure at the Large Hadron Collider : Experimental Review. 2018.

[12] J. R. Andersen et al. Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report. In *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*, 2016.

[13] Simone Marzani, Gregory Soyez, and Michael Spannowsky. Looking inside jets: an introduction to jet substructure and boosted-object phenomenology. 2019. [Lect. Notes Phys.958,pp.(2019)].

[14] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*, D48:3160–3166, 1993.

[15] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better jet clustering algorithms. *JHEP*, 08:001, 1997.

[16] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[17] Jeff Tseng and Hannah Evans. Sequential recombination algorithm for jet clustering and background subtraction. *Phys. Rev.*, D88:014044, 2013.

[18] Gavin P. Salam. Towards Jetography. *Eur. Phys. J.*, C67:637–686, 2010.

[19] Andrew J. Larkoski, Simone Marzani, Gregory Soyez, and Jesse Thaler. Soft Drop. *JHEP*, 05:146, 2014.

[20] Mrinal Dasgupta, Alessandro Fregoso, Simone Marzani, and Gavin P. Salam. Towards an understanding of jet substructure. *JHEP*, 09:029, 2013.

[21] M. Dasgupta and G. P. Salam. Resummation of nonglobal QCD observables. *Phys. Lett.*, B512:323–330, 2001.

[22] Mrinal Dasgupta, Alessandro Fregoso, Simone Marzani, and Alexander Powling. Jet substructure with analytical methods. *Eur. Phys. J.*, C73(11):2623, 2013.

[23] Mrinal Dasgupta, Alexander Powling, and Andrzej Siodmok. On jet substructure methods for signal jets. *JHEP*, 08:079, 2015.

[24] Andrew J. Larkoski, Ian Moult, and Duff Neill. Non-Global Logarithms, Factorization, and the Soft Substructure of Jets. *JHEP*, 09:143, 2015.

[25] Mrinal Dasgupta, Kamel Khelifa-Kerfa, Simone Marzani, and Michael Spannowsky. On jet mass distributions in Z+jet and dijet processes at the LHC. *JHEP*, 10:126, 2012.

[26] Yang-Ting Chien, Randall Kelley, Matthew D. Schwartz, and Hua Xing Zhu. Resummation of Jet Mass at Hadron Colliders. *Phys. Rev.*, D87(1):014010, 2013.

[27] Teppo T. Jouttenus, Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. Jet mass spectra in Higgs boson plus one jet at next-to-next-to-leading logarithmic order. *Phys. Rev.*, D88(5):054031, 2013.

[28] Leandro G. Almeida, Stephen D. Ellis, Christopher Lee, George Sterman, Ilmo Sung, and Jonathan R. Walsh. Comparing and counting logs in direct and effective methods of QCD resummation. *JHEP*, 04:174, 2014.

[29] Ze Long Liu, Chong Sheng Li, Jian Wang, and Yan Wang. Resummation prediction on the jet mass spectrum in one-jet inclusive production at the LHC. *JHEP*, 04:005, 2015.

[30] Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. Dissecting Soft Radiation with Factorization. *Phys. Rev. Lett.*, 114(9):092001, 2015.

[31] Kamel Khelifa-Kerfa and Yazid Delenda. Non-global logarithms at finite N_c beyond leading order. *JHEP*, 03:094, 2015.

[32] Christopher Frye, Andrew J. Larkoski, Matthew D. Schwartz, and Kai Yan. Factorization for groomed jet substructure beyond the next-to-leading logarithm. 2016.

[33] Daniel W. Kolodrubetz, Piotr Pietrulewicz, Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. Factorization for Jet Radius Logarithms in Jet Mass Spectra at the LHC. 2016.

[34] John M. Campbell and R. Keith Ellis. An Update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60:113006, 1999.

[35] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 07:018, 2011.

[36] The CMS collaboration. Measurements of the differential jet cross section as a function of the jet mass in dijet events from proton-proton collisions at $\sqrt{s} = 13$ tev. *Journal of High Energy Physics*, 2018(11):113, Nov 2018.

[37] Morad Aaboud et al. Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 121:092001, Aug 2018.

[38] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), Jul 2014.

[39] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J.*, C76(3):155, 2016.

[40] Albert M Sirunyan et al. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. 2019.

[41] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.

[42] CMS Collaboration. Particle–flow event reconstruction in CMS and performance for jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$. *CMS Physics Analysis Summary*, CMS-PAS-PFT-09-001, 2009.

[43] CMS Collaboration. Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector. *CMS Physics Analysis Summary*, CMS-PAS-PFT-10-001, 2010.

[44] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[45] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur.Phys.J.*, C72:1896, 2012.

[46] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett.*, B659:119–126, 2008.

[47] Adrian Cameron and Pravin Trivedi. *Microeconometrics : methods and applications*. Cambridge University Press, New York, 2005.

[48] CMS Collaboration. Search for massive resonances decaying into pairs of boosted W and Z bosons at $\sqrt{s}$ = 13 TeV. 2015.

[49] CMS Collaboration. CMS Luminosity Measurement for the 2015 Data Taking Period. CMS Physics Analysis Summary CMS-PAS-LUM-15-001, 2016.

[50] Stefan Schmitt. TUnfold: an algorithm for correcting migration effects in high energy physics. *JINST*, 7:T10003, 2012.

[51] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth.*, A362:487–498, 1995.