Fachbereich Mathematik und Naturwissenschaften
Fachgruppe Physik
Bergische Universität Wuppertal

Resonance Searches with the $t\bar{t}$ Invariant Mass Distribution

measured with the DØ Experiment at $\sqrt{s} = 1.96\,\mathrm{TeV}$

Thorsten D. Schliephake

Diese Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20100804-114229-1
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20100804-114229-1]

For my parents and my grandmother

Contents

CHAPTER 1

Introduction

Understanding the universe, its birth and its future is one of the biggest motivations in physics. In order to understand the cosmos, the fundamental particles forming the universe, the components our matter is built of need to be known and understood. Over time physicists have built a theory which describes the physics of the known fundamental particles very well: the Standard Model (SM) of particle physics [1, 2, 3, 4, 5, 6].

The SM describes the particles, their interactions and phenomena with high precision. So far no proven deviations from the SM have been found, though recently evidence for possible physics beyond the SM has been observed [7]. The SM is not describing the mass of the elementary particles however and even with the addition of the Higgs mechanism giving mass to the particles, we have no full theory for all four fundamental forces. We know the model needs to be extended or replaced by another one, as gravitation is not included in the SM.

Having a theory which describes all fundamental particles found so far and all but one fundamental interaction is a great success. However, all this describes about 4 % of the universe we live in. 23 % is dark matter and 73 % is dark energy. Dark matter is believed to interact only through gravity and maybe the weak force, which makes it hardly observable. Dark energy is even more elusive. Among other theories the cosmologic constant and scalar fields are discussed to describe it. One should also note that other models exist which for example modify the Newtonian law of gravity.

The Higgs mechanism has become the most popular model for mass generation. Alternative theories like Super Symmetry (SUSY), large Extra Dimensions, Technicolor, String Theory, to name just a few, have spread to describe the necessary

mass generation or new particles. As proof for new physics beyond the SM has not been found yet, one assumes that new physics will manifest itself at a larger energy scale and therefore a higher particle mass. Particles with high masses are therefore presumed to be a window to test the SM for deviations caused by new physics.

The heaviest fundamental particle which is in our reach is the top quark. Its mass is almost as large as that of a complete tungsten atom. It is so heavy, that it decays faster than it can hadronize. It seems the perfect probe to study new physics at the moment.

In this analysis the top quark is used as a probe to search for a new resonance, whose properties are similar to a SM Z boson but is much more massive. This analysis will study $t\bar{t}$ decays to search for an excess in the invariant mass distribution of the $t\bar{t}$ pairs.

Resonant states are suggested for massive Z-like bosons in extended gauge theories [8], Kaluza Klein states of the gluon or Z [9, 10], axigluons [11], topcolor [12], and other beyond the Standard Model theories. Independent of the exact model a resonant production mechanism should be visible in the $t\bar{t}$ invariant mass distribution.

In this thesis a model-independent search for a narrow-width heavy resonance X decaying into $t\bar{t}$ is performed. In the SM, the top quark decays into a W boson and a b quark nearly 100% of the time, which has been proven experimentally, too [13]. The $t\bar{t}$ event signature is fully determined by the W boson decay modes. In this analysis, only the lepton+jets final state, which results from the leptonic decay of one of the W bosons and the hadronic decay of the other, is considered. The event signature is an isolated electron or muon with high transverse momentum, large transverse energy imbalance due to the undetected neutrino, and at least three jets, two of which result from the hadronization of b quarks.

The analyzed dataset corresponds to a combined integrated luminosity of $3.6\,\mathrm{fb}^{-1}$ in the e+jets and in the μ+jets channel. This collision data was collected between August 2002 and July 2008.

The analysis uses data with three or more jets to select $t\bar{t}$ events. The signal-to-background ratio is improved by identifying b-jets using a neural network based b-tagging algorithm. After b-tagging, the dominant physics background for a resonance signal is non-resonant SM $t\bar{t}$ production. Smaller contributions arise from the direct production of W bosons in association with three or more jets, as well as instrumental background originating from multijet processes with jets faking isolated leptons. The search for resonant production is performed by examining the reconstructed $t\bar{t}$ invariant mass distribution.

A topcolor assisted extended technicolor model [14] was used as a reference

to quote mass limits for Z' bosons. In this model, a large top quark mass can be generated through the formation of a dynamical $t\bar{t}$ condensate, Z', which is formed by a new strong gauge force coupling preferentially to the third generation of fermions. The Z' couples weakly and symmetrically to the first and second generation and strongly only to the third generation of quarks and has no couplings to leptons. This model results in a predicted cross section for $t\bar{t}$ production larger than the SM prediction.

This analysis has previously been performed in Run I by DØ [15] and CDF [16] as well as in Run II by DØ [17, 18] and CDF [19, 20, 21, 22]. So far neither experiment found any evidence for a $t\bar{t}$ resonance. Limits obtained on $\sigma_X \times B(X \to t\bar{t})$ were used to quote a limit on the mass of such a topcolor Z'.

Studies made in this analysis have shown that the limit on $\sigma_X \times B(X \to t\bar{t})$ is valid for any narrow resonance with vector or axial vector couplings mixed in any combination. This is because $\sigma_{t\bar{t}}$ is rather insensitive to the difference of vector or axial vector couplings. Therefore a model independent limit on $\sigma_X \times B(X \to t\bar{t})$ is set, depending on the resonance mass.

CHAPTER 2

Theory

The understanding of physics has made great progress in the last century. Starting with Planck around the turn of the 20$^{\text{th}}$ century, particle physics has been evolved to a set of complex theories, each describing one part of nature. By now, we believe to describe the universe by four fundamental interactions:

- The *gravitational force* is the vital one on large scales, as in astrophysics.
- The *electromagnetic force* is important for all charged objects, mainly known from the interactions between atoms.
- The *weak force* is about $^1/_{1000}$ of the scale of a nucleus or atom, most commonly known in radioactive decays.
- The *strong force* keeps the nucleus and even its components, the hadrons, bound by gluing the quarks together.

The weak and the electromagnetic force have been unified in an electroweak force already. It is believed that this force can be unified in a Grand Unified Theory (GUT) at the GUT-(energy)-scale with the strong force. So far, no description of a quantized gravity exists. Physicists have not come up yet with a good theory which describes at the same time gravity on large and small scales such as quantum theory.

This chapter starts with a description of the standard model and will later on focus on a single particle of the SM, the top quark. Finally, since it is important for the analysis process of this thesis the introduction will end in a selection of possible particles which are not predicted by the SM and can give rise to new physics.

2.1 The Standard Model

This section gives an introduction to the theories known as the Standard Model of particle physics. Extensive testing of the theoretical predictions given by the SM describes nature to a very precise level, hence the model became the Standard Model.

The model is extremely precise up to energies of several hundred GeV although not predicting masses and not including gravity. This leads to a stunning success of a model which is known to be incomplete. As a more detailed explanation of the SM can be found in various textbooks [1, 2, 3, 4, 5, 6] only a short summary is given here.

At this stage it should clearly be noted, that throughout this thesis units appropriate for particle physics will be used. This means the two fundamental constants $\hbar$ and c will be set to one meaning $\hbar = c = 1$ unless explicitly stated otherwise. Energy, momentum and mass are measured in units of GeV, while length and time have units of GeV^{-1}. Macroscopic units like the ones used in measuring detector dimensions will still be given in their appropriate SI units.

The SM began to develop in the late 1960's when Glashow, Salam, and Weinberg merged the weak and electromagnetic interactions into a single theory [23]. At this time the quark hypothesis was developed which gave the large number of new particles found in the 1950's a building scheme like the periodic table in chemistry. Those new particles interact with a new force coupling to a charge called color.

The complete SM is based on the symmetry groups $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ where C denotes color, L the weak isospin and Y the hypercharge. The different parts will be explained in more detail in the following sections.

2.1.1 Bosons & Fermions

The SM describes a number of fermions, which can be identified as the fundamental building particles of matter, and gauge bosons that carry the force. Table 2.1 lists the quarks and some of their quantum numbers while table 2.2 does the same for the leptons. The gauge bosons are listed in table 2.3.

2.1.2 Strong Interaction

The quark hypothesis was derived by Gell-Mann, Ne'eman, and Zweig in the 1960's [26, 27, 28]. The current quark model describes particles with a quantum

Generation	Quark Flavor	Spin [$\hbar$]	Charge [e]	Mass [GeV]
I	up (u)	$1/2$	$2/3$	$0.0015 - 0.003$
	down (d)	$1/2$	$-1/3$	$0.0035 - 0.006$
II	charm (c)	$1/2$	$2/3$	$1.27^{+0.07}_{-0.11}$
	strange (s)	$1/2$	$-1/3$	$0.104^{+0.026}_{-0.034}$
III	top (t)	$1/2$	$2/3$	173.1 ± 1.6
	bottom (b)	$1/2$	$-1/3$	$4.20^{+0.17}_{-0.07}$

Table 2.1: List of quarks in the SM with their spin, charge and mass. The u, d, and s quark masses are estimates of so called "current-quark masses" in a mass-independent subtraction scheme such as $\overline{\text{MS}}$ at a scale $\mu \sim 2\,\text{GeV}$. The c and b quark masses are the "running masses" in the $\overline{\text{MS}}$ scheme [24]. The top quark mass is taken from the most recent combined CDF and DØ measurement and refers to the pole mass [25].

Generation	Lepton Flavor	Spin [$\hbar$]	Charge [e]	Mass [MeV]
I	electron (e)	$1/2$	-1	0.510998910(13)
	electron neutrino (ν_e)	$1/2$	0	$< 2 \cdot 10^{-6}$
II	muon (μ)	$1/2$	-1	105.658367(4)
	muon neutrino (ν_μ)	$1/2$	0	< 0.19
III	tau (τ)	$1/2$	-1	1776.84(17)
	tau neutrino (ν_τ)	$1/2$	0	< 18.2

Table 2.2: List of leptons in the SM with their spin, charge and mass [24].

Force	Gauge Boson	Spin [$\hbar$]	Charge [e]	Mass [GeV]
gravitation	graviton (?)	2	0	0
electromagnetic	photon (γ)	1	0	0
weak	$W^\pm$	1	± 1	80.398(25)
	Z	1	0	91.1876(21)
strong	8 gluons (g_i with $i = 1, ..., 8$)	1	0	0

Table 2.3: List of bosons in the SM with their spin, charge and mass [24].

number which was later identified with the charge of the strong force: the color charge. The charge is similar to the electric charge with the differnce of bearing three different charges called red, blue and green. These three different charges were necessary to describe baryons like the Δ^{++}. Being a fermion, it needs to have an antisymmetric wave function. It consists of three u quarks which are totally symmetric in space, spin and flavor. Hence a forth quantum number, color, is needed.

The corresponding field theory is called quantum chromodynamics (QCD) based on the $\mathrm{SU}(3)_C$ group mentioned above. The group has eight generators, leading to the eight gauge bosons: the gluons, which are massless. The gluons carry a color charge, more precise a color and anticolor charge, which gives rise to a self coupling. The self coupling of the gluons leads to an increasing coupling strength with increasing distance. Eventually the energy in the field gets high enough to produce (anti-)particles and break one color field into two smaller ones. The field energy gained is used for the creation of the new (anti-)particles.

In nature, no color charged objects but only color neutral objects, so called color singlets, are observed. This is also known as confinement. To be even more precise, only objects which are invariant under rotation in $\mathrm{SU}(3)_C$ can be observed. For example the π^0 meson is given by $\sqrt{1/3}(R\bar{R} + G\bar{G} + B\bar{B})$ instead of a single $R\bar{R}$ state similar to the normalization of the u and d quarks is given by $\sqrt{1/2}(u\bar{u} + d\bar{d})$. Therefore any color has to be canceled by its corresponding anti-color or must be existent in equal of all three colors forming a colorless or "white" state.

This leads to two basic classes of hadrons: mesons and baryons. The mesons are built of $q\bar{q}$ pairs with color and corresponding anti-color. Baryons consist of qqq or $\bar{q}\bar{q}\bar{q}$ triplets, where each color is present in equal amount. Though other more complex objects like penta-quarks ($qqqq\bar{q}$) could be imagined, none have been observed so far.

The constant force of the strong interaction over long distances results in an increasing potential in the force field. At small distances, the field behaves like a Coulomb potential. At small distances, translating into small wavelength or high energies of the probing particle, the colored objects behave similar to unbound objects, hence the energy in the force field is small. This is called asymptotic freedom. The small coupling at very high energies allows the use of perturbation theory. The coupling parameter of the strong force α_s is energy dependent. The perturbation theory which is used to describe quantum chromodynamics is only valid down to a certain energy. Then other models need to be applied to describe the physical process. A model referred to as hadronization or fragmentation describes how color charged particles transfer to colorless compound objects like baryons or mesons. The high energy used in hadron colliders makes it possible to

describe the proton/(anti-)proton interaction as a collision by individual partons from each of the (anti-)protons.

2.1.3 Electroweak Interaction & SM Higgs Mechanism

The first SM interaction described by a gauge theory was the electromagnetic interaction with only one gauge boson, the photon. The symmetry group in which quantum electrodynamics (QED) is described is $U(1)$. The invariance under local gauge transformation is equivalent to the photon coupling to electromagnetic charges. The perturbation is done in the coupling constant α. The coupling α rises with increasing energy or decreasing distance. At energies around 1 GeV, α is about $1/137$ and allows perturbative expansion. Another neat feature is that α rises slowly with the transferred momentum Q^2 and therefore allows corrections up to high energy levels.

The weak interaction can be described in the same way as QCD and QED, but the charge that the gauge bosons couples to is the weak isospin. The gauge bosons W^i are massive and couple only to left handed particles. This orders the left handed particles in isospin doublets and the right handed ones in singlets, as can be seen in table 2.4.

Particles				Isospin I	3^{rd} Component I_3	Hypercharge Y	Charge $Q[e]$
leptons	$\binom{\nu_{e_L}}{e_L}$	$\binom{\nu_{\mu_L}}{\mu_L}$	$\binom{\nu_{\tau_L}}{\tau_L}$	$\binom{1/2}{1/2}$	$\binom{1/2}{-1/2}$	$\binom{-1/2}{-1/2}$	$\binom{0}{-1}$
	e_R	μ_R	τ_R	0	0	-1	-1
quarks	$\binom{u_L}{d'_L}$	$\binom{c_L}{s'_L}$	$\binom{t_L}{b'_L}$	$\binom{1/2}{1/2}$	$\binom{1/2}{-1/2}$	$\binom{1/6}{1/6}$	$\binom{2/3}{-1/3}$
	u_R	c_R	t_R	0	0	$2/3$	$2/3$
	d_R	s_R	b_R	0	0	$-1/3$	$-1/3$

Table 2.4: Isospin doublets and singlets for quarks and leptons.

The electromagnetic force and the weak force are unified in the SM and can be described based on a $\mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ symmetry group. This theory is called GWS-theory after Glashow, Weinberg and Salam [6, 23, 29]. The massless gauge bosons W^1, W^2, and W^3 from $SU(2)_L$ and the B from the $U(1)_Y$ group do not commute with each other due to the fact that $SU(2)_L$ is not an Abelian group. The observed gauge bosons $W^\pm$, Z, and γ can be represented as a linear combination of the fundamental gauge bosons:

$$W^\pm = (W^1 \mp W^2)/\sqrt{2} \tag{2.1}$$

$$Z = W^3 \cos\theta_W - B^0 \sin\theta_W \tag{2.2}$$

$$\gamma = W^3 \sin\theta_W + B^0 \cos\theta_W \tag{2.3}$$

where θ_W is the weak mixing angle determined by the coupling constants g_W for the $SU(2)_L$ group and g'_W for the $U(1)$ group, given as:

$$\sin\theta_W = \frac{g'_W}{\sqrt{g_W^2 + g_W'^2}}. \tag{2.4}$$

The gauge bosons are massless at first as the simple introduction of mass terms by hand would violate local gauge invariance. The massive gauge bosons can be explained by introducing a spontaneous symmetry breaking by the Higgs mechanism [30]. The introduction of a complex scalar isospin doublet

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi^+ \\ \phi^0\end{pmatrix} \tag{2.5}$$

serves this purpose and to be consistent with the gauge symmetry, the potential needs to have the form

$$V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{2.6}$$

with $\mu^2 < 0$ and $\lambda > 0$. The vacuum expectation of this field is given by [1]:

$$\phi_0 = \sqrt{\frac{1}{2}}\begin{pmatrix}0 \\ v\end{pmatrix} \tag{2.7}$$

where

$$v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{2.8}$$

While v can experimentally be determined, the parameter λ cannot. This leaves the value of μ^2 undetermined, which translates into the mass of the scalar Higgs boson through:

$$M_h = \sqrt{-2\mu^2}. \tag{2.9}$$

The Higgs boson mass is therefore not predicted and the Higgs boson particle has not yet been found by experimental searches. The mass of the gauge bosons can be written as:

$$m_W^2 = \frac{e^2v^2}{4\sin^2\theta_W} \qquad \text{and} \qquad m_Z^2 = \frac{e^2v^2}{4\cos^2\theta_W\sin^2\theta_W}. \tag{2.10}$$

They can either be used to determine the mixing angle:

$$\sin\theta_W = 1 - \frac{m_W^2}{m_Z^2} \tag{2.11}$$

or the measurement of the mixing angle leads to a coupling of the masses of the gauge bosons, which tests the SM.

Besides the weak gauge bosons, the fermions are also not massless. The mass of a lepton can be introduced by the Yukawa coupling of leptons with the Higgs field. The resulting mass term can be written as:

$$M_l = \frac{G_l v}{\sqrt{2}}. \tag{2.12}$$

The size of the Yukawa coupling constant G_l is unknown and has to be evaluated by measuring the mass of the leptons. Similarly, quarks acquire their masses from their Yukawa coupling to the Higgs field. The mass eigenstates of the quarks (d, s, b) are not identical to the weak eigenstates (d', s', b'). The relation is a unitary transformation [31]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = U_d \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} = U_u \begin{pmatrix} u \\ c \\ t \end{pmatrix}. \tag{2.13}$$

Usually the matrix U_u is chosen to be the unity matrix, to make the up-type quark weak and mass eigenstates to be identical. The matrix $V = U_u^\dagger U_d$ is called Cabbibo-Kobayashi-Maskawa (CKM) Matrix and can be written as [24]:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.14}$$

$$= \begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.999133^{+0.000044}_{-0.000043} \end{pmatrix}.$$

The probability of a transition from a quark q_i into a quark q_j by emitting a (virtual) W boson is proportional to the corresponding CKM-Matrix element squared: $|V_{q_i q_j}|^2$. The unitarity requirement leaves nine free parameters five of which can be absorbed in the quark phases. This leaves four free parameters in the matrix, three of which are mixing angles and one is a phase, that causes the CP violation in the SM. The diagonal elements turn out to be close to unity while the transfers to the non-neighboring family are suppressed by the squared order of the transfer to the next family.

Neutrinos are not massless. The electroweak and mass eigenstates differ as well. The transformation matrix is similar to the CKM-Matrix and is called

Maki-Nakagawa-Sakata (MNS) matrix [32]. It connects the electroweak eigenstates (ν_1,ν_2,ν_3) to the flavor eigenstates (ν_e,ν_μ,ν_τ). This is the reason why neutrino oscillation can be observed. However, in most practical cases neutrinos are considered as massless.

2.2 Top Quarks

The top quark, the heaviest known SM particle, was discovered in 1995 by CDF and DØ [33, 34]. It completes the table of quarks and was predicted by theory since the discovery of the b quark in 1977. Electroweak fits and indirect measurements made before 1995 by the SLD and LEP experiments pointed to a high mass. With its current measurement of 173.1 GeV [25], its mass is as high as a heavy metallic atom like tungsten. Compared to other SM particles this is about twice the mass of the massive weak vector bosons or 40 times the mass of the next heaviest quark, namely the b quark. The very high mass makes it likely, that new beyond SM theory processes can be tested. The Yukawa coupling is close to unity which also might have consequences for the electroweak symmetry breaking. Though theoretically well known and described, only a few properties of the top quark were actually measured experimentally. A more complete overview of top quark physics can be found in [35, 36].

The high mass, which makes it so distinct, results in a very short lifetime. The average lifetime of a top quark of $4.2 \cdot 10^{-25}$ s is much lower than the time needed for typical hadronization processes, which is about 10^{-23} s. The top quark therefore decays before it can form a bound quark state giving the opportunity to observe a *free* quark. This is also the reason why the mass measurement can be very precise as there are no truly compound objects before the decay.

In a $X \rightarrow t\bar{t}$ search, the SM $t\bar{t}$ pair production is an irreducible background. The production and decay of the top quarks will therefore be described shortly. In the following, no explicit distinction between particles and their antiparticles is made unless explicitly noted.

2.2.1 Top Quark Production at Hadron Colliders

Due to its high mass top quarks can currently be produced at two accelerators in the world: the Tevatron accelerator and the now operational and running LHC accelerator. Top quark pairs can be produced via quark-antiquark ($q\bar{q}$) annihilation or by gluon gluon (gg) fusion. The leading order Feynman diagrams for both processes are shown in figure 2.1. The theoretical cross section at the Tevatron used for this analysis was calculated by Langenfeld, Moch, and Uwer

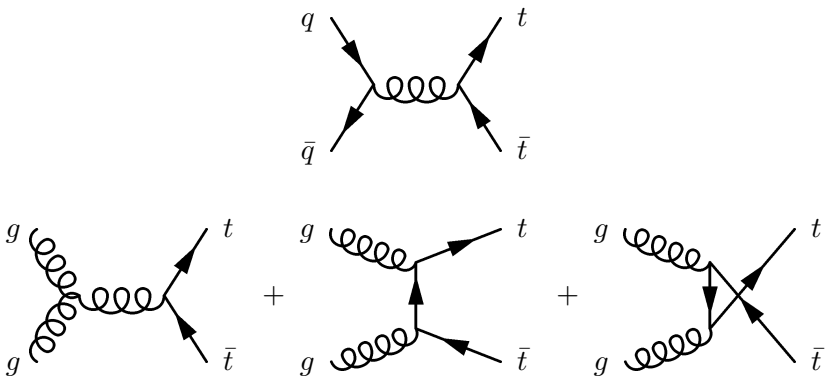


Figure 2.1: Leading order Feynman diagrams for quark-antiquark annihilation and gluon gluon fusion.

to be $7.34^{+0.23}_{-0.38}$ pb [37].

In compound objects, each of the partons inside carries a part of the total momentum. In protons, those partons consist of (anti-)quarks and gluons. The probability that a certain type of parton carries a certain fraction x of the total momentum can be parametrized in a parton distribution function (PDF), which is displayed in figure 2.2. These functions are measured in deep inelastic scattering to observe different momentum fractions and hard scattering scales Q^2. Valence quarks carry a rather high amount of the total momentum. Therefore the curves for the up and down quarks rise significantly over those of the sea quarks. One can also see that there are roughly two times as much up quarks than there are down quarks. For the low momentum fraction, the gluons are most likely to be observed.

At the Tevatron as a $p\bar{p}$ collider the quark-antiquark production is favored and the amount of gluon gluon fusion is less than 15 %. The high center of mass energy of the LHC and the fact that it is a pp collider makes it more likely to produce $t\bar{t}$ via gluon gluon fusion. This can be understood by calculating the needed momentum fraction of each of the partons inside the (anti-)proton. The Tevatron collider needs on average $x \geq 0.09$, while the LHC needs (at 14 TeV) on average $x = 0.0125$ for each parton. By looking at the parton distribution functions, one can calculate the percentage of partons fulfilling the required x. Additionally, the hard scattering process scale Q^2 needs to be known, as the PDFs are a function of x and Q^2. For $t\bar{t}$ pair production this scale is most often set to $Q^2 = m_t^2$. At the Tevatron collider this leads to $\sim 85\,\%$ $q\bar{q}$-annihilation and $\sim 15\,\%$ gg-fusion, while at the LHC over 90 % are from gg fusion. This change due to the large difference in the needed x value for the partons.

Besides pair production the production of a single top quark with a vector boson is also possible and has recently been observed and measured at the Tevatron [39, 40, 41]. At the Tevatron only the t-channel and the s-channel processes have been

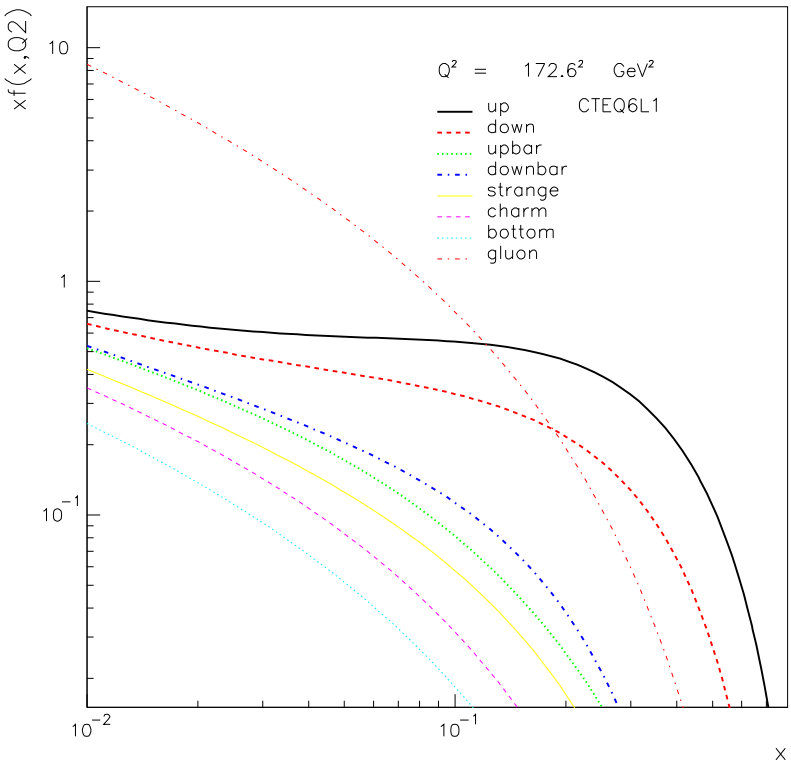


Figure 2.2: Parton distribution function of the proton [38] with CTEQ6L1 parametrization.

verified. The combined cross section measurement from CDF and DØ yielded a cross section of $2.76^{+0.58}_{-0.47}$ pb. The leading order single top quark production Feynman diagrams are displayed in figure 2.3. As can be seen, each diagram has a Wtb vertex, which can also be used for the measurement of the CKM matrix element V_{tb} .

2.2.2 Top Quark Decay

The top quarks decay almost completely into a W boson and a b quark. Therefore the decay of the top quark is determined by the decay of the W boson, which either decays into a lepton and a corresponding neutrino or into a quark-antiquark pair. With two top quarks in an event, this leads to several possible final states listed in figure 2.4 each having its unique structure:

- In the **alljets** channel, both of the W bosons decay into quark-antiquark

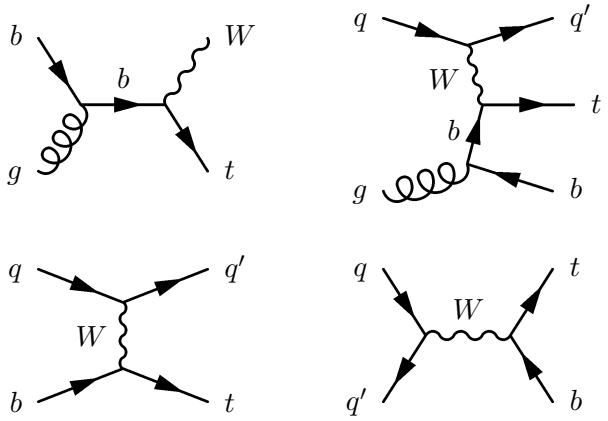


Figure 2.3: Feynman diagrams for the electroweak single top production at hadron colliders.

pairs. This leads to six jets from (anti-)quarks, two of which are b quark jets. This channel has the highest branching ratio of about 46 % but suffers from the high multijet background.

- In the **lepton + jets** channel (l+jets) the W bosons decay semi-leptonically, which means one W bosons decays into two quarks and the other into a lepton neutrino pair. This channel has a detectable charged lepton, missing transverse energy from the neutrino, and four jets, two of which are b quark jets. This channel has about 30 % branching ratio taking into account the τ decaying into electrons and muons. The lepton + jets channel will therefore only include electrons and muons as well as the contribution from the τs decaying into electrons or muons.

- In the **dileptonic** channel both W bosons decay into a lepton and a neutrino, leading to two leptons, two b quark jets and missing transverse energy in the detector. Leptons here are electrons and muons as well as τs decaying into electrons and muons. The branching ratio including the τ contribution is about 7 %. This channel suffers from low statistics and the ambiguity, i.e. missing momentum vectors, given by two indetectable neutrinos, but has a small background.

- In the channels containing τ leptons, decaying into hadrons or explicit requirements for τ decays, the final state is more complicated and of no interest in this analysis. They are therefore neglegted in the following.

This analysis concentrates on the semi-leptonic (l+jets) decay channel, which still has good statistics and a detectable lepton to suppress multijet background.

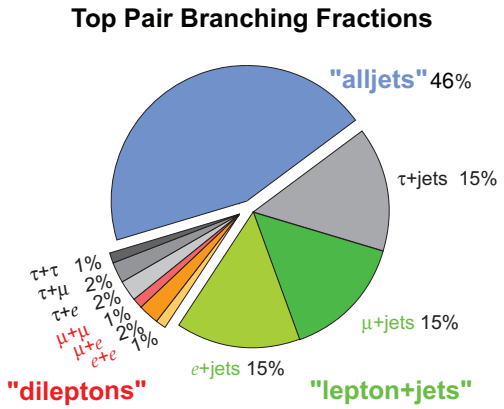


Figure 2.4: $t\bar{t}$ decay channels and the corresponding branching ratios. The branching ratios correspond to the theoretical predictions at tree level [42].

2.2.3 Connection to the SM Higgs Boson

Since the mass of the top quark is very high, its Yukawa coupling constant is close to unity. It plays a vital role in higher order corrections, where loops are dominated by top quarks and Higgs boson contributions. In tree level approximation the expectation value v and the coupling constant g can be determined by the fine structure constant α and the Fermi coupling G_F. The W boson mass can then be written as:

$$m_W^2 = \frac{1}{2} m_Z^2 \left(1 + \sqrt{1 - \frac{4\pi\alpha}{\sqrt{2} G_F m_Z^2}} \right) = \frac{\pi\alpha}{\sqrt{2}} G_F \sin^2 \theta_W . \tag{2.15}$$

The contribution from virtual top quarks and higgs bosons leads to a correction term Δr, which modifies 2.15 to:

$$m_W^2 = \frac{\pi\alpha}{\sqrt{2} G_F \sin^2 \theta_W (1 - \Delta r)} . \tag{2.16}$$

where the contributions of the virtual top quark loops are given by:

$$(\Delta r)_t \sim -\frac{3 G_F}{8\sqrt{2}\pi^2 \tan^2 \theta_W} \cdot m_t^2 \tag{2.17}$$

and for virtual Higgs boson loops by:

$$(\Delta r)_h \sim -\frac{11 G_F m_Z^2 \cos^2 \theta_W}{24\sqrt{2}\pi^2} \cdot ln \frac{m_h^2}{m_Z^2} . \tag{2.18}$$

This leads to an indirect constraint on the Higgs mass as shown in figure 2.5. The resulting fit to the electroweak data can be seen in figure 2.6, where the left yellow exclusion was set during the LEP era, while the right yellow strip was set by a recent combined CDF and DØ measurement, where the exclusion region for the Higgs boson mass is between 162 and 166 GeV [43].

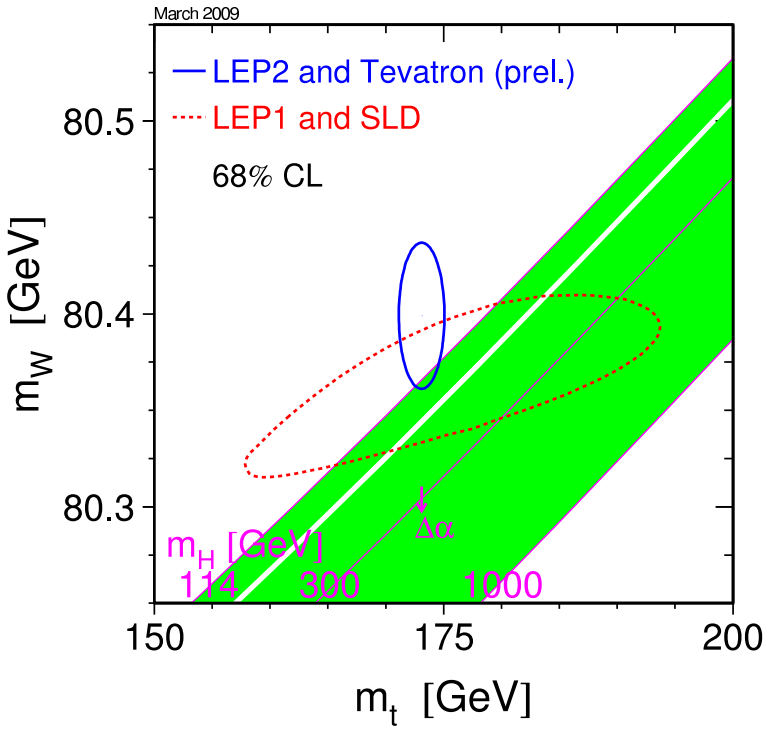


Figure 2.5: Measurement of the top and W mass (m_t, m_W) with a global fit to the electroweak data. The correlation between m_t and m_W as expected in the Standard Model is also shown for different Higgs boson masses. The possible range of the SM Higgs boson mass is colored in green [44].

2.3 $t\bar{t}$ Resonances

In the SM no bound $t\bar{t}$ states are expected, as the lifetime of the top quark is too short. The following sections will summarize a number of possibilities to observe resonant $t\bar{t}$ production. In general, the resonances presented in the following sections must be neutral bosons and have at least twice the mass of a top quark. A selection taken from [45] of possible resonance types is given in table 2.5.

In gg fusion processes, a spin zero resonance could give a peak-dip structure due to the interference term in the matrix element. The Tevatron has a negligible

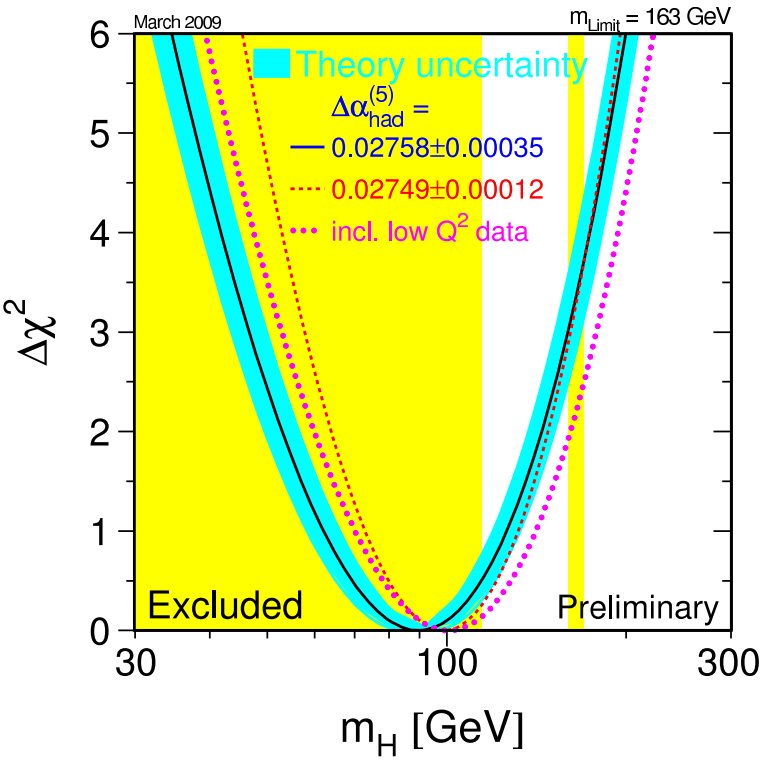


Figure 2.6: Indirect measurement of the Higgs boson mass with the lower limit at 95 % confidence level on the Higgs boson mass. The yellow area shows the excluded Higgs boson masses [44].

amount of $gg \to X \to t\bar{t}$. Any such resonance would be buried in the $q\bar{q} \to t\bar{t}$ continuum. A spin one resonance would not interfere with the SM $t\bar{t}$ production, as this is forbidden by Yang's theorem [56, 57, 58]. Therefore it is sufficient to search for a peak structure in the $t\bar{t}$ invariant mass distribution. At the LHC, it would be worthwhile searching for a peak-dip structure.

The presented analysis covers particles with spin one and color singlet quantum numbers. The vector and axial vector particles are studied without a specific underlying theory. This is done to broaden the validity of the result to general small width spin one color singlet resonances. As benchmark model, the leptophobic topcolor assisted technicolor model described below is used to quote mass exclusion limits. This model is therefore described in more detail in contrast to the other models. A broad overview of Z' gauge bosons can be found in [59].

Spin	Colour	Parity $(1, \gamma_5)$	Examples or Ref.
0	0	(1,0)	SM, MSSM, 2HDM; Ref. [46, 47, 48]
0	0	(0,1)	MSSM, 2HDM; Ref. [47, 48]
0	8	(1,0), (1,1)	Lepto-Quarks, Techni-π^0; Ref. [49, 50]
0	8	(1,0), (0,1)	Techni-π^0; Ref. [49, 50]
1	0	(SM,SM)	Z'
1	0	(1,0)	vector
1	0	(0,1)	axial vector
1	0	(1,1)	vector-left
1	0	(1,-1)	vector-right
1	8	(1,0)	coloron, KK gluon; Ref. [51, 52, 53]
1	8	(0,1)	axigluon; Ref. [52]
2	0	–	graviton; Ref. [54, 55]

Table 2.5: Beyond SM particles decaying into $t\bar{t}$ quark pairs [45]. Resonances covered in this analysis are color singlet and have spin one.

2.3.1 SM Higgs

The only possibility to observe resonant $t\bar{t}$ production in the SM is a heavy Higgs boson decaying into a $t\bar{t}$ pair. As the Higgs boson couples to the mass of the fermions, the top quark is the perfect probe to study a SM Higgs boson. Assuming a Higgs boson with the mass of about 400 GeV, the branching ratio into a $t\bar{t}$ pair is about 14 % [60]. The heavy Higgs production through gluon fusion at the Tevatron is of the order of 0.1 pb [61]. This makes such a resonance practically impossible to observe. Additionally, the electroweak fits presented in the last section favor a light Higgs boson.

2.3.2 Higgs Doublet Models (2HDM, MHDM)

The SM Higgs boson suffers from a problem which is known as the fine tuning problem [62, 63]. The mass of the Higgs boson shows strong quantum corrections from bosons and fermions it couples to. While electroweak fits set the physical mass to the order of ~ 100 GeV the quantum corrections are in the order of the cut off scale. This is usually chosen to be the Planck scale $\Lambda_{\text{Planck}} \simeq 10^{19}$ GeV where new physics should appear if the model is complete. Choosing the Planck scale would make one loop corrections to the squared Higgs mass 30 magnitudes larger than the physical Higgs mass. This seems unnatural.

One way to solve this difficulty is to introduce a super symmetry theory (SUSY), where super-symmetrical partners to fermions, which are bosons, and to bosons, which are fermions, are introduced. As the coupling has opposite sign

for fermions and bosons, this would cancel out the corrections.

The simplest modification, without the need for super symmetry, would be the introduction of an additional complex Higgs doublet, i.e. four new degrees of freedom, which form four additional Higgs bosons [64]. The advantage of the mass generation with the two Higgs doublet fields remain similar to the SM Higgs theory, removing three degrees for the mass generation of the weak gauge bosons. The five Higgs bosons emerging are massive through self interaction. The bosons are a pair of neutral CP-even scalars h and H, where the first one is similar to the SM Higgs boson, a neutral CP-odd scalar A, and a pair of charged Higgs H^+ and H^-. The neutral bosons h, H, and A could decay into a $t\bar{t}$ quark pair.

In principal, this can be extended by any number of additional Higgs doublets adding more complex fields. These multi Higgs doublet models (MHDM) [65, 66] have $2n-1$ neutral and $2(n-1)$ charged observable scalars.

Exclusion limits with more than the SM Higgs boson have a lower bound on the mass of the lightest Higgs boson. Neutral Higgs bosons in a minimal supersymmetric standard model (MSSM) are excluded up to masses of $m_h < 85$ GeV and $m_A < 86$ GeV, where $m_h < m_H$ [67].

2.3.3 Axigluons & Colorons

Another way of introducing mass is extending the QCD symmetry group $SU(3)$ to $SU(3) \times SU(3)$ [68, 52], broken at a certain energy scale.

For the axigluon model, the $SU(3)_{L+R}$ is extended to $SU(3)_L \times SU(3)_R$ [69, 70], which leads to an octet of massless gluons g and massive axigluons A. They have an axial vector coupling $1/2 g_s \gamma_\mu \gamma_5 \lambda^a$, where g_s is the usual strong coupling and λ^a are the Gell-Mann matrices.

The universal coloron model [71] gives rise to a class of models with extended color like topcolor and technicolor models described below. In general the symmetry is broken around the TeV-scale to a $SU(3)_I$ and $SU(3)_{II}$ group with different couplings g_1 and g_2, where $g_2 \ll g_1$. The massive colorons couple to quarks through a $1/2 g_s \cot(g) \gamma_\mu \lambda^a$, where g is the mixing angle given by $\cot g = g_1/g_2$.

In both theories eight massless gauge bosons emerge, which are the gluons from QCD. The remaining eight bosons are an octet of color charged heavy particles. Depending on the used symmetry group and couplings to quarks, the additional particles are either vector bosons (Colorons) or axial vector bosons (Axigluons). Both interact by the strong coupling with quarks and should therefore be visible through their decay into top quarks. Limits obtained at the Tevatron are listed below in table 2.6.

Particle	Coupling	Excluded Mass Range [GeV]
Axigluon	-	$m_A < 910$
Coloron	$\cot\theta = 1$	$m_C < 800$, $895 < m_C < 1960$
	$\cot\theta = 1.5$	$150 < m_C < 950$
	$\cot\theta = 2$	$m_C < 955$, $1030 < m_C < 3200$

Table 2.6: Coloron and axigluon exclusion limits at the Tevatron [52] at 95 % C.L. each.

2.3.4 Topcolor Assisted Technicolor Z'

One of the theories beyond the SM tries to give a possible explanation for the mass hierarchy observed in the three generations by introducing a new charge called technicolor (TC) [72, 73, 74]. This provides a dynamical description of the electromagnetic symmetry breaking. Technicolor achieves this by introducing a new force analogous to the strong force. This model predicts a corresponding techni-particle for each SM particle. The electroweak symmetry breaking takes place through the condensation of the techni-particles at the technicolor scale of $\Lambda_{TC} \sim 1$ TeV.

Technicolor alone does not address the problem of flavor. It does not explain multiple generations, which is why extended technicolor (ETC) was developed. In ETC, ordinary $SU(3)$ color and $SU(N)_{TC}$ technicolor and flavor symmetries are unified into the ETC gauge group G_{ETC}. The energy scale of ETC is very high compared to TC, which is about 0.1 - 1 TeV. The ETC scale is

$$\Lambda_{ETC} \sim \frac{14 \text{ TeV}}{N^{3/4}}, \tag{2.19}$$

where N is the number of techni-doublets.

As flavor changing neutral currents are not observed, the coupling of this extended technicolor model $\alpha_{TC}(\mu)$ has to develop slowly over the range of $\lambda_{TC} \lesssim \mu \lesssim M_{ETC}$, where M_{ETC} is the mass of a typical gauge boson. This model can evolve naturally all masses except for the high top mass.

Another model called topcolor can produce high top masses. This involves a dynamical $t\bar{t}$ condensate at the scale Λ_t generated by a new strong gauge force, which couples strongly to the third generation with a coupling parameter g_1 and weakly to the first and second generation with a coupling parameter g_2. To preserve electroweak $SU(2)$, topcolor needs to treat t_L and b_L in the same way. In order to prevent a large b quark mass, the weak isospin must be violated and t_R and b_R treated differently. In order to simulate the SM in low energy regions and especially the small violation of the weak isospin, the scale must be very high: $\Lambda_t \sim 10^{15}\,\text{GeV} \gg m_t$. This would require fine tuning of the couplings to $\Lambda^2/m_t^2 \simeq 10^{25}$, which seems unnatural.

Therefore a combined topcolor assisted technicolor (TC2) theory is assumed, where the electroweak symmetry breaking is mainly driven by the technicolor interactions, which are strong near 1 TeV. All light quarks, leptons, and technipions are generated by ETC and the topcolor will generate the $t\bar{t}$ condensate and the high top mass. Both TC and topcolor have their scale around 1 TeV in this scenario, while the ETC scale is around 100 TeV in order to suppress flavor changing neutral currents (FCNC). This also requires a walking coupling, where the coupling constant is slowly rising with energy scale.

In this scheme, a separate $SU(3)$ for color and weak hypercharge $U(1)$ exists for the third $SU(3)_1 \times U(1)_1$ and the first two generations $SU(3)_2 \times U(1)_2$. The third generation has a strong coupling at the scale of about 1 TeV, while the first two have a weak coupling. This symmetry breaking leads to the formation of massive gauge bosons: a color octet of "colorons" V_s and a color singlet Z'. Top condensation is forced by the fact that $SU(3)_1 \times U(1)_1$ is supercritical for top quarks but subcritical for bottom quarks.

The Z' particles predicted by the TC2 theories still differ from each other. One model predicts a Z', which has, additionally to the above features, no coupling to leptons. This leads to a preferred decay into $t\bar{t}$ quark pairs, which increases the cross section for the $t\bar{t}$ quark production. Table 2.7 lists the cross sections for this model at the Tevatron for the CETQ6L1 PDFs and a width of 1.2 % of the resonance mass.

$m_{Z'}$[GeV]	$\sigma(p\bar{p} \to Z' \to t\bar{t})$[pb]
360	10.00
400	12.23
450	8.61
500	5.44
550	3.27
600	2.01
650	1.23
700	0.76
750	0.45
800	0.28
850	0.17
900	0.10
950	0.06
1000	0.04

Table 2.7: Theoretically predicted $\sigma_X \times B(X - t\bar{t})$ for $Z' \to t\bar{t}$ with a width of $\Gamma_Z = 0.012 m_{Z'}$ calculated for the CTEQ6L1 PDFs.

2.3.5 Kaluza Klein Particles

String theory predicts at least six extra dimensions beyond the known four of space-time. Those dimensions are believed to have periodic boundary conditions like small loops and have the size of the inverse Planck scale, which would leave them unobservable. Other theories predict so called large extra dimensions in the order of $\sim$ TeV^{-1}, which can be probed with current hadron colliders.

The d'Alembert operator in d dimensions can be split in the usual four dimensional d'Alembert operator plus the second derivatives in the additional dimensions. The summed momenta in the extra dimension plus the true rest mass squared form the visible mass in four dimensions given in the Klein-Gordon equation by:

$$\left(\Box^{(4)} \underbrace{- \sum_{\mu=5}^{d} \frac{n_\mu^2}{R^2} - m^2}_{-m^2_{\text{visible}}} \right) \Psi = 0 \,. \tag{2.20}$$

As the extra dimensions have periodic boundary conditions, the momenta in those are quantized similar to a harmonic oscillator. The ground state describes a SM particle with mass m.

For one extra dimension, the excitations will have an equidistant increase in mass. These excitations are then called Kaluza-Klein (KK) towers. More dimensions will allow additional resonances between the ones found in one dimension. This leads to a denser spectrum at high energies up to a quasi continuum in the limit of very high energies. Additionally, the extra dimensions may be of different size or warped.

As the rest mass of the particle is added to the mass acquired by the excitation, massless particles would be the first observed at low energies. With one extra dimension the first KK excitations of SM gauge fields ($W^\pm$, Z, γ and g) lie above $\simeq$ 4 TeV [10]. The second excitation would be beyond the reach of the LHC even at several times design luminosity.

Searches for $t\bar{t}$ resonances have been performed at the Tevatron. For a KK gluon of the bulk Randall-Sundrum model [75], a lower mass limit of 800 GeV was observed.

CHAPTER 3

Experimental Setup

The Fermi National Accelerator Laboratory (FNAL) is located east of Chicago. The FNAL complex houses a variety of accelerators, the biggest one being the Tevatron. There are two experiments located around collision points of the Tevatron as well as a couple of fixed target experiments which are operated from smaller accelerators. Additionally neutrino experiments like MiniBooNE are located on the site. A large theoretical staff complements the experimental one. The Tevatron provides collisions of protons and antiprotons with an energy of 1.96 TeV at two collision points, which are B0 and D0. At these collision points two multipurpose detectors are located. These are the Collider Detector at Fermilab (CDF) and DØ . A description of the latter will be provided in the following sections, as the observed data forms the foundation of this analysis.

3.1 Tevatron

The following sections are dedicated to the Tevatron, its preaccelerators, as well as the performance of the accelerator and its integrated luminosity.

3.1.1 Accelerator Chain

The Tevatron is the last stage in a chain of accelerators shown in figure 3.1. Beam generation starts with ionizing Hydrogen to H^- and accelerating it to 750 keV with an electrostatic Cockcroft-Walton accelerator. The H^- ions are accelerated in a 130 m long linear accelerator (LINAC) to 400 MeV and then the electrons

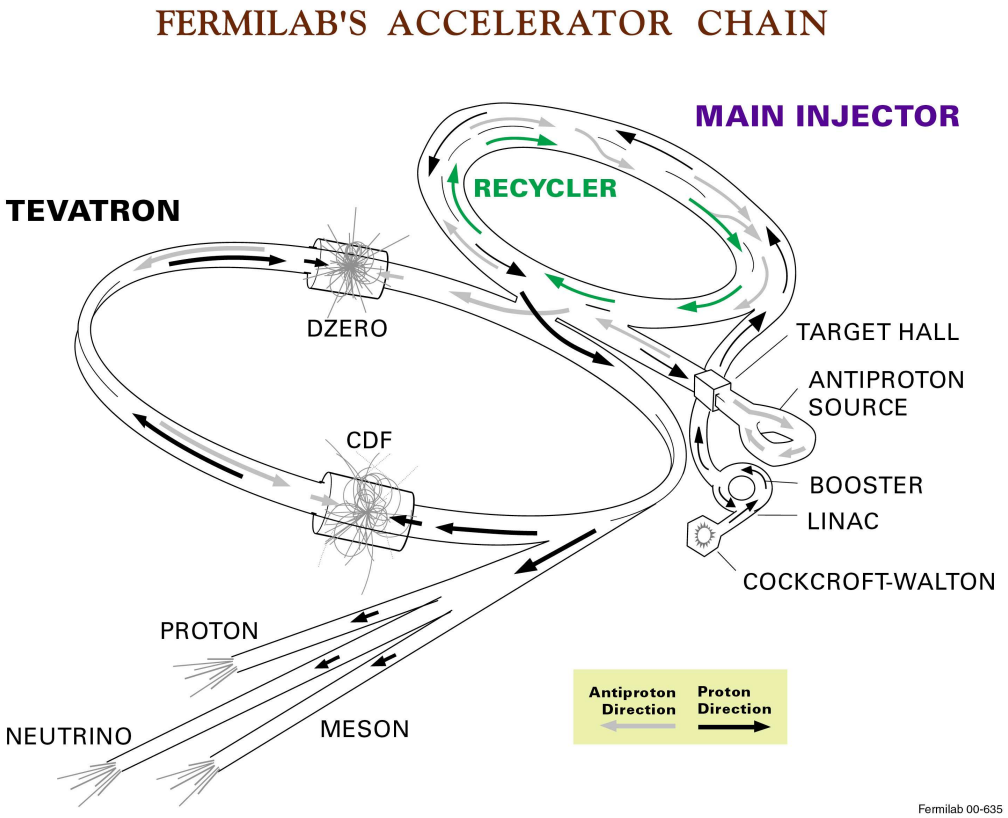


Figure 3.1: Accelerator chain [76].

are stripped off by a thin carbon foil to get bare protons. These are transferred to the Booster, which is 75 m in radius, and accelerated to 8 GeV.

After around 20.000 revolutions the protons are transferred to the Main Injector, which is about 7 times bigger in radius than the booster. For antiproton production, which is continuously taking place, the protons are accelerated to 120 GeV and shot onto a nickel target with copper cooling discs. For about 1 million protons shot onto the target, about 200 antiprotons are produced. They are filtered with a magnetic field produced by a lithium lens, a cylindrical lithium conductor operated at 500 kA, producing a field gradiant of 740 T/m. The antiprotons are then transferred to the debuncher and accumulator, where they undergo stochastic and electron cooling and are then stored at an energy of 8 GeV.

A period during which the protons collide with the antiprotons is called store. When enough antiprotons have been accumulated, a process which takes about 20 hours, the old store is dropped and the shot setup takes place. At first, protons are injected into the main ring to fill 36 bunches. After all proton bunches have

successfully been filled into the Tevatron, the antiprotons are loaded in another 36 bunches as has already been done with the protons. The 36 bunches are subdivided into three trains spaced by 7µs each, containing 12 bunches separated by 396 ns. Each proton bunch contains $\mathcal{O}(10^{11})$ protons, while the antiproton bunches each contain $\mathcal{O}(10^{10})$ antiprotons. The two beams are then accelerated from 150 GeV to 980 GeV and brought to collision at the two interaction points called B0 and D0. The detector at the former point is known as CDF.

Since 2006, the recycler located along the ceiling of the main injector is in use, where the antiprotons are transferred to in regular intervals from the accumulator. This is done because the accumulation of antiprotons is more effective, hence faster, with less antiprotons being in the accumulator. As well the recycler is used for electron cooling of the antiprotons.

3.1.2 Performance of the Tevatron

The main parameter, the performance of an accelerator, is the luminosity $\mathcal{L}$. The instantaneous luminosity is given by:

$$\mathcal{L} = n \cdot f \cdot \frac{N_{\mathrm{p}} N_{\bar{\mathrm{p}}}}{2\pi\sigma_{\mathrm{p}}\sigma_{\bar{\mathrm{p}}}} \cdot F \ , \tag{3.1}$$

where n is the number of bunches, f the revolution frequency, and N_{p} ($N_{\bar{\mathrm{p}}}$) the number of (anti-)protons in each bunch. Besides π [77] the denominator contains the spatial width of the (anti-)proton bunches σ_{p} ($\sigma_{\bar{\mathrm{p}}}$). F is a form factor, which depends on several parameters like for example the length of the bunches. $\mathcal{L}$ is measured in $cm^{-2}s^{-1}$, which is a particle flux.

While the initial luminosity for a store was as low as $\mathcal{L} = 25 \cdot 10^{30} cm^{-2}s^{-1}$ in Run I the instantaneous Run II luminosity has reached more than $\mathcal{L} = 4.024 \cdot 10^{32} cm^{-2}s^{-1}$ on the 16^{th} of April 2010.

Integrating the instantaneous luminosity over time, one gets the integrated luminosity $\mathcal{L}_{\mathrm{int}}$ measured in inverse barn ($1\,b = 10^{-28}\,\mathrm{m}^2$). This makes it easy to calculate the expected number of events N of a certain physic process given by:

$$N = \sigma \cdot \mathcal{L}_{\mathrm{int}} \ , \tag{3.2}$$

where σ is the cross section of the process of interest.

The integrated luminosity delivered so far by the Tevatron and the recorded luminosity by DØ is shown in figure 3.2. The data taking efficiency is about 90 % for DØ . The baseline goal for the Tevatron Run II was 4.4 fb^{-1}, which has clearly been reached, though the design goal of 8.5 fb^{-1} until August 2009 was missed

by either 1.5 fb^{-1} or 8 months. Of the currently 7.81 fb^{-1} recorded by DØ this anlysis uses 3.6 fb^{-1}.

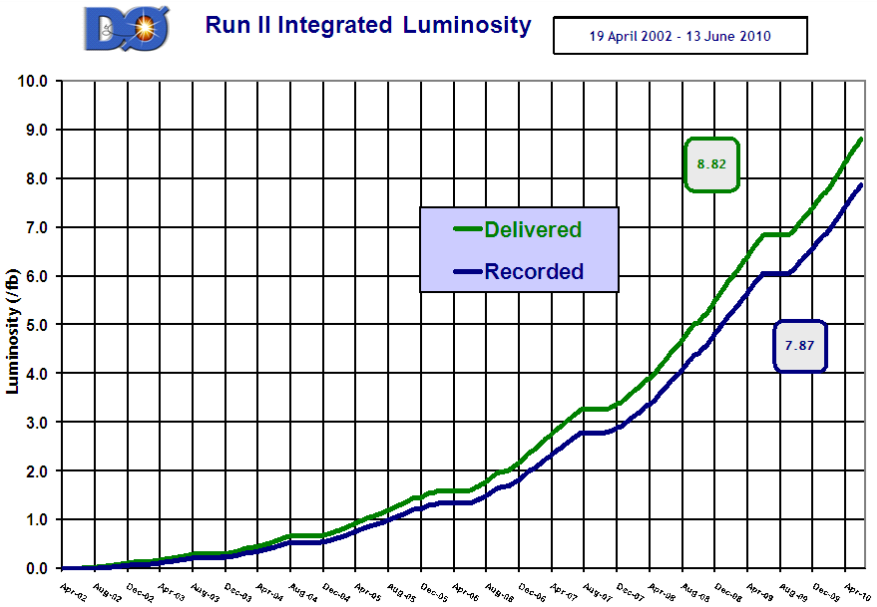


Figure 3.2: Delivered and recorded Luminosity [78].

3.2 The DØ-Detector

The DØ-Detector is a 4π multipurpose detector which combines precision charged tracking with fast projective calorimetry and fine grained muon detection. A side view of the detector is shown in figure 3.3. It extends 20 m in length and 13 m in height. The main components are the tracking system, the calorimeter and the muon system. A brief overview of the detector is given in the following sections. A complete description of the DØ-Detector can be found in [79, 80, 81].

The detector coordinate system is right handed. The positive z-axis points along the proton beam, the positive x-axis points horizontally in the Tevatron ring, and the positive y-axis points upwards. The transverse plane is defined by the x- and y-axis. A cylindrical coordinate system is used with:

$$r = \sqrt{x^2 + y^2}, \qquad \phi = \arctan\tfrac{x}{y}, \qquad \eta = -\ln\tan(\tfrac{\theta}{2}), \tag{3.3}$$

where θ denotes the azimuthal angle, and η the pseudo-rapidity, which is equal to the rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{3.4}$$

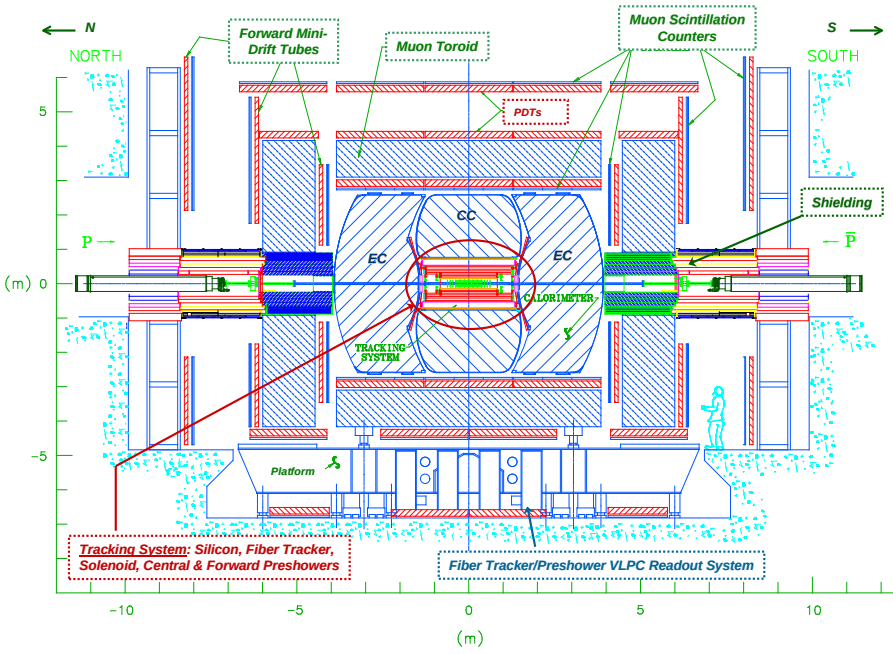


Figure 3.3: Side view of the DØ detector [82].

for massless particles.

At DØ, two definitions for η exist. One with respect to the center of the detector, called η_{det}. The other, the physics η is defined relative to the primary vertex of the event. The latter will be used in the following, unless stated otherwise.

3.2.1 Tracking System

The tracking system is the innermost part of the detector. It consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT) around the SMT. Both are encased in a solenoid magnet with a field of 2 T. The length of the super conducting magnet is 2.73 m including the cryostats, which can be seen in figure 3.4. The magnetic lines can be seen in figure 3.5.

The cryostat and magnet together yield 1.1 radiation lengths, X_0, where one radiation length X_0 is both the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung and $7/9$ of the mean free path for pair production by a high-energy photon. The tracking system has a momentum resolution of:

$$\Delta p_T / \text{GeV} = 0.002 (p_T / \text{GeV})^2. \tag{3.5}$$

This resolution is much better than the one of the muon system as can be seen in section 3.2.3. Therefore, it is also important for measuring muon momenta.

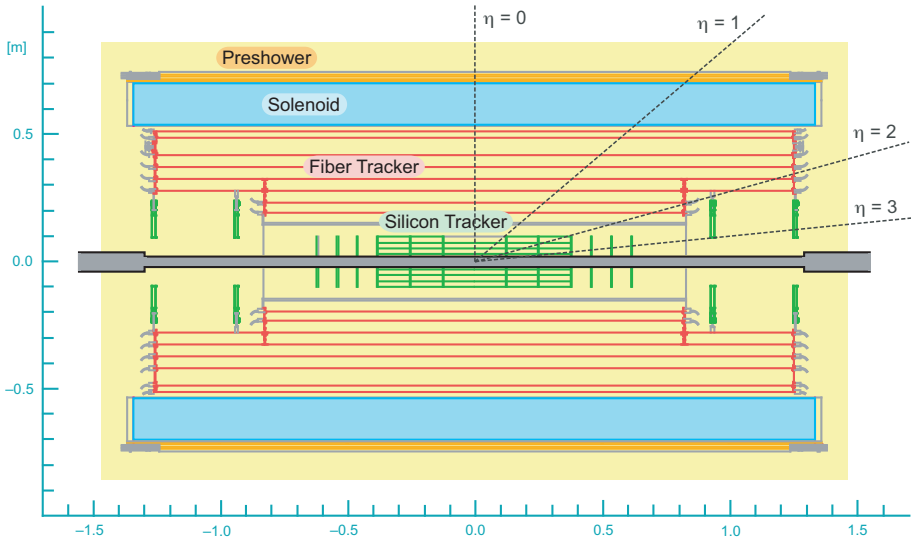


Figure 3.4: Schematic picture of the central tracking system [83].

The Silicon Microstrip Tracker

The SMT is the part of the detector nearest to the interaction point. Its purpose is the tracking and determination of primary and secondary vertices. Secondary vertices are especially important for the identification of b-jets (as in section 4.4.2). The SMT consists of six barrels and 16 discs as can be seen in figure 3.6. The barrels are used for measuring the r-ϕ coordinate and for the transverse momentum of the track, while the discs gather additional information about $r-z$, giving a three-dimensional track information.

Each barrel is made of four layers of silicon readout modules, which are called ladders. The first and second layer consist of 12 ladders each, while the third and fourth layer contain 24 ladders. Each layer has two sublayers. The central four barrels use double-sided double metal (DSDM) detectors in layer 1 and 3. Layer 2 and 4 use single-sided (SS) technology in the central barrels. In the outermost barrels double-sided (DS) detectors are used. The DSDM detectors consist of axial and 90° stereo strips, while the DS detectors consist of axial strips and ones with a stereo angle of 2°. SS detectors only have axial strips.

The discs that intersperse the barrel are called F-discs, which consist of 12 double-sided wedge-detectors. In the forward region, the so called H-discs are placed for η up to ± 3. The H-discs contain 24 wedges each, which are made of two back-to-back single-sided wedges. All discs are planar modules. The double sided H-discs have an effective stereo angle of 30°. Two single sided H wedges taken together form a double sided sensor with a stereo angle of 15°. Each of the four sublayers consist of two sublayers overlapping each other to cover the

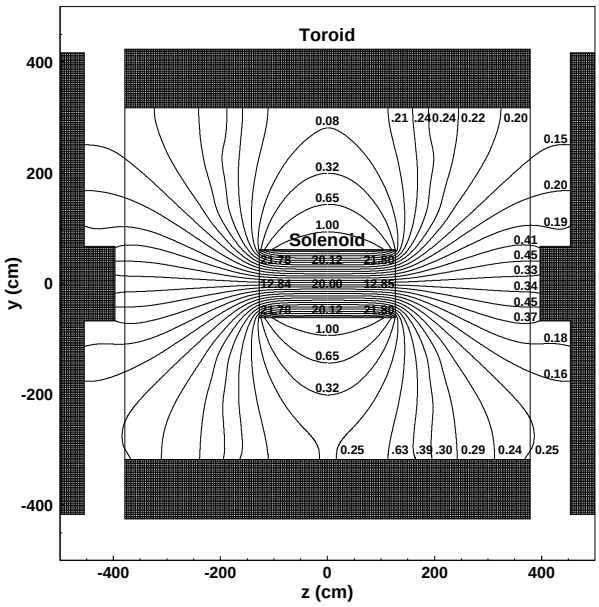


Figure 3.5: Illustration of the magnetic field inside the DØ detector [81].

complete parameter space in ϕ.

In 2005, a new layer 0 was introduced into the SMT [84] closest to the interaction point. By including layer 0, a better vertex resolution could be achieved.

Central Fiber Tracker

The central fiber tracker (CFT) is built around the SMT as can bee seen in figure 3.7. The CFT covers a range of $|\eta| < 1.62$ and consists of 76800 scintillating fibers, which are mounted on eight concentric cylinders. The two innermost cylinders have a length of 1.66 m, while the six outer ones are 2.52 m in length each. Each cylinder consists of a double layer of fibers, one layer extending in axial direction (axial layer) and a second layer with a stereo angle in ϕ of $\pm 3°$ (stereo layer). The stereo layers in the first, third, fifth, and seventh cylinder have a $+3°$ and the rest a $-3°$ orientation.

The scintillating fibers are coupled to clear fiber waveguides. These conduct light to visible light photon counters (VLPC). At this point the light signal is converted into an electric signal and finally read out. The diameter of a fiber is 835 μm and it is either 1.66 m or 2.53 m long. Only one end is connected to a waveguide, while the other end is covered with sputtered aluminum coating (reflectivity 90%) to reflect the light signal.

A typical charged particle produces about ten γs per fiber. The VLPC quantum

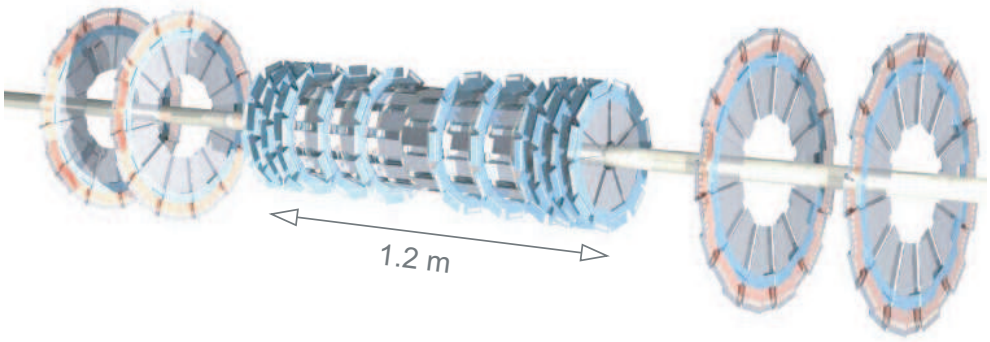


Figure 3.6: The design of discs and barrels in the SMT [81].

efficiency of $\geq 75\%$ and a high gain amplifies each photon into 2200 to 65000 electrons.

3.2.2 Calorimeter and Shower Detectors

This section covers the calorimeter and shower detectors. The preshower detector which gives an increased resolution for the already showered objects from the inner layer of the detector is covered first. Following are the barrel and endcap electromagnetic and hadronic calorimeters, as well as the intercryostat detectors, which cover the region between the barrel and the endcap.

The Preshower Detector

The preshower detector is arranged around the solenoid magnet. It consists of a central preshower (CPS) which covers the region of $|\eta| < 1.3$ and a forward preshower (FPS) covering $1.5 \leq |\eta| \leq 2.5$. Its purpose is an increase in electron and photon identification efficiency and a better background rejection.

The CPS is located around a 5.6 mm lead radiator ($\sim 1X_0$). Together with the $0.9\,X_0$ radiation length solenoid, the CPS forms almost two radiation lengths of material in radial direction. Depending on the angle of the entering particle this extends to up to $4X_0$. The CPS has three layers of scintillator strips, all of which consist of 1280 strips each. Every strip is read out by two wavelength-shifting fibers.

The FPS is located at the two ends of the calorimeter. The two FPS on the north and south side of the detector are mounted on the calorimeter cryostats. Both detectors consist of two layers of double planes of scintillators separated by

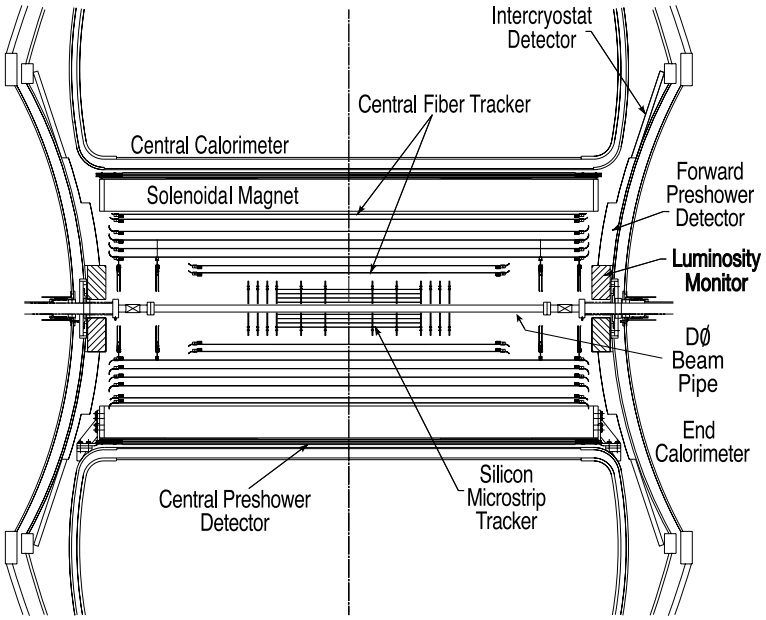


Figure 3.7: Cross-sectional view of the SMT in the barrel region [81].

11 mm ($2X_0$) thick lead-stainless-steel absorbers. In the innermost layer, passing charged particles deposit minimum ionizing signals. This part therefore is called minimum ionizing particle (MIP) layer. The outermost layer is called shower layer as the charged particles and photons cause shower signals here.

Calorimeter

The main purpose of the calorimeter is to identify and measure the kinetic energy of particles. The calorimeter consists of three parts: the central calorimeter (CC) as well as the end cap calorimeters (EC) on the north (ECN) and the south side (ECS) of the detector as shown in figure 3.8.

The EC and CC are sampling calorimeters with absorber plates made of uranium surrounded by liquid argon, which is kept at 90 K. Each calorimeter is mounted in a separate cryostat to maintain the low temperature. Figure 3.9 shows the coverage of the calorimeter as well as the segmentation of the absorber plates.

The CC covers $|\eta| < 1$, while ECN and ECS cover up to $|\eta| \sim 4$. All of the three calorimeters consist of four electromagnetic layers (EM). These layers are surrounded by three fine and one coarse hadronic layer. The absorber plates used in the EM are thin depleted uranium plates, which are 3 mm thick for the CC and 4 mm for the ECs. The fine hadronic layer uses thin uranium-niobium

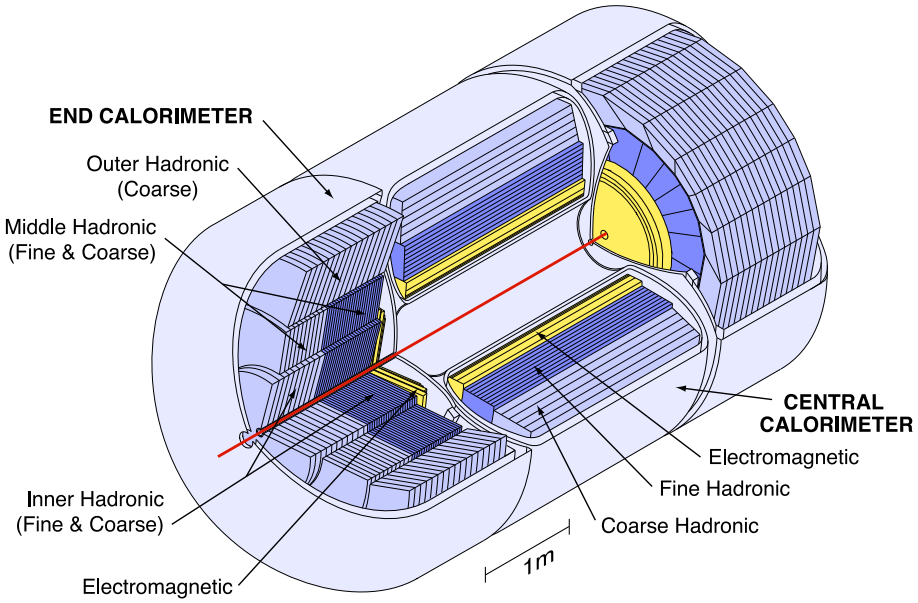


Figure 3.8: Overview of the DØ liquid argon calorimeter [81].

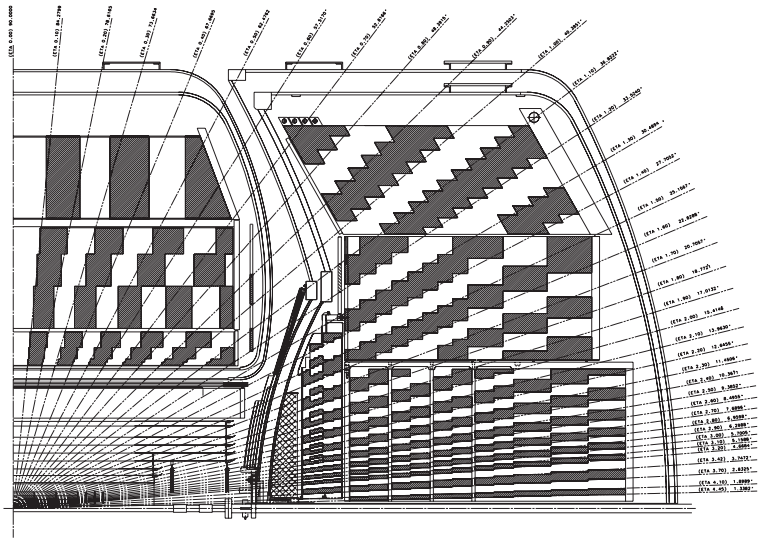


Figure 3.9: Schematic view of a quadrant of the DØ calorimeter, showing segmentation and as shaded areas merged cells for readout [81].

alloy, while the coarse hadronic uses 46.5 mm thick copper plates in the CC and stainless steel plates in the ECs.

The segmentation of the fine and coarse hadronic calorimeter is 0.1×0.1 in the $\eta - \phi$ plane, as well as in the first, second, and fourth layer of the EM. The third layer, which is at the shower maximum, has a segmentation which is twice as fine to give a more precise determination of the EM shower centroids. In the central region the electromagnetic calorimeter has $20.6\,X_0$ and the hadronic one has $\sim 6.4\,\lambda_A$ [85, 86], where λ_A is the nuclear interaction length. This is the distance at which in average the energy of a hadron has decreased to $1/e$ of its original value.

Electromagnetic showers are completely absorbed in the EM calorimeter, while hadronic particles start to shower in the EM, but the main energy deposition is in the fine and coarse hadronic calorimeter. This differing shower behavior can be used to distinguish the electromagnetic particles from the hadronic ones.

The energy of electromagnetic and hadronic particles can be parametrized as [85, 86]:

$$\frac{\Delta E}{E} = \sqrt{\frac{S^2}{E} + \frac{N^2}{E^2} + C^2}, \tag{3.6}$$

where N denotes the instrumental effects like uranium noise, pedestal subtraction and S denotes fluctuations in the energy deposition of the particles. C takes calibration uncertainties into account. An overview of the parameters for DØ is given in table 3.1. In contrast to the tracking system, the calorimeter resolution gets better with increasing energy.

Object	C	S $[\sqrt{\text{GeV}}]$	N [GeV]
Electrons, Photons	0.041	0.15	0.29
Jets	0.036	1.05	2.13

Table 3.1: Energy resolution parameters [87, 88].

Inter-Cryostat Detector

The inter-cryostat detector (ICR) fills the gap between $0.8 < |\eta| < 1.4$ to provide better resolution in this otherwise sparsely covered region. Scintillating tiles attached to the surface of the EC fill the region of $1.1 < |\eta| < 1.4$. Standard calorimeter readout cells have been added within the CC and EC cryostats to fill the remaining gaps, which are called massless gaps. The ICD consists of 16

scintillating tiles each covering $\Delta\eta \times \Delta\Phi \sim 0.3 \times 0.4$, where every single tile consists of 12 subtiles covering $\Delta\eta \times \Delta\Phi \sim 0.1 \times 0.1$.

3.2.3 Muon System

The muon system is the outermost layer of the DØ-Detector. Its purpose is to identify and measure muons. A magnet with a toroidal field, which has a strength of 1.8 T, provides an independent measurement of the muons. The field lines are presented in figure 3.5.

The muon system consists of two parts, the wide angle muon system (WAMUS) or central muon system and the forward muon system (FAMUS). The central system covers the range in pseudorapidity of $|\eta| \leq 1$ while the forward system ranges in $1 \leq |\eta| \leq 2$. Both systems consist of three layers called A-, B-, and C-layer.

In the central system, all three layers consist of proportional drift tubes (PDTs) and scintillating counters. A single PDT module is made of an aluminum tube and is $2.8 \times 5.2\,\mathrm{m}^2$ in size.

Layer A is arranged directly around the calorimeter followed by the toroidal magnet and by the layers B and C. Around 90% of the instrumented space is covered by at least two layers of the muon system. Around $\sim$ 55% is even covered by all three layers.

Each of the three layers includes three decks of drift cells, except for layer A which is consisting of four decks. The bottom layer A has three decks. A typical drift chamber contains 72 (3 decks) or 96 (4 decks) cells. Each cell is about 10.1 cm wide. The electronics to read out the wires are located at the end of each drift chamber.

On the top, side, and bottom of the C-layer scintillation counters are installed. Those at the bottom are called bottom counters where the rest is called cosmic gap. Both counters provide a fast timing signal, which enables the association of the muons to a certain bunch crossing as well as discrimination from cosmic ray backgrounds.

Furthermore, layer A's PDTs are covered with so called $A\phi$ scintillation counters which give a fast detector response to reject out-of-time backscattering and allows the triggering and identification of muons. Their timing resolution is about 2 ns.

In the forward system, mini drift tubes (MDTs) are used instead of the PDTs. The complete system consists of the end cap toroidal magnets, the MDTs of the A-, B-, and C-layer, three layers of scintillators and the beam pipe shielding.

Layer A is housed inside the toroid similar to the central system and consists of 4 planes of MDTs which are mounted parallel to the magnetic field lines. Layers B and C are located outside of the toroid consisting of three planes each. All drift tube have eight cells, each of which covers an area of about $9.4 \times 9.4\,\mathrm{mm}^2$.

Trigger scintillation counters designed to provide good timing resolution are added outside of each layer. They consist of eight parts in ϕ, whereas each part contains 96 counters. Figure 3.10 shows a view of the PDTs and MDTs.

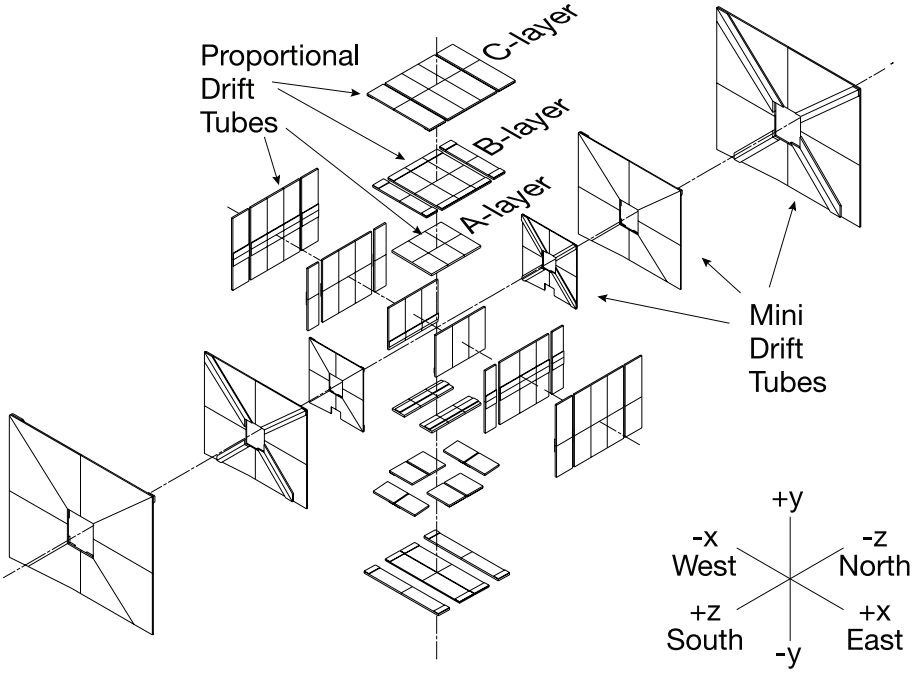


Figure 3.10: Exploded view drawing of the PDTs and MDTs in the muon system [81].

Figure 3.11 shows the position of the cosmic gap, the bottom counters, the Aϕ scintillators, and the trigger scintillation counters.

A toroidal magnet has been installed into the muon system not only for having a standalone muon-system momentum-measurement but also for having a high field strength for low p_T cutoffs in the level 1 muon trigger as well as getting a better matching of the muon with the central tracks. Also π and K decays can be rejected better. Furthermore the momentum resolution for muons with high p_T is highly increased. The polarity of the magnet is reversed periodically during data taking runs to minimize systematic uncertainties due to asymmetries in the detector.

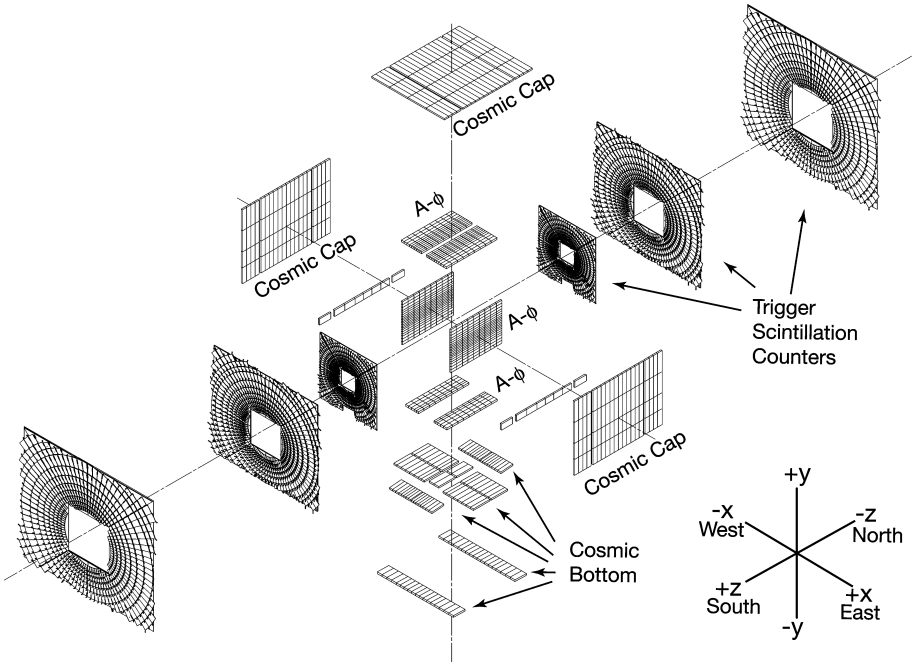


Figure 3.11: Exploded view drawing of the scintillation detectors of the muon system [81].

3.2.4 Trigger and Data Acquisition

At the Tevatron, collisions take place every 396 ns at each interaction point, which leads to an average collision rate of 25 MHz during the trains and an overall rate of 1.7 MHz taking the abort gaps into account. Current computing technology and funding limit the rate of events, which can be stored and reconstructed to the order of 50-100 Hz.

The majority of the events are inelastic proton antiproton scattering, which are less interesting than, for example, top quark events. Measurements taken to reduce the number of events by preselecting the physically most interesting events, are called triggers [81, 89, 90].

DØ uses a three level trigger system to reduce the number of events in steps and to gain more time at each level for the examination of the event. Level 1 (L1) is a pure hardware built trigger, using signals from each detector separately. L1 has an accept rate of about 2 kHz. Level 2 (L2) consists of hardware engines embedded in microprocessors and reduces the rate to about 1 kHz. They can already use multiple detector components and complex objects. Level 3 (L3) is a farm of microprocessors, where on each processor a full event is reconstructed. This farm has an output rate of about 50-100 Hz. Events passing L3 are recorded for offline reconstruction. An overview can be found in figure 3.12.

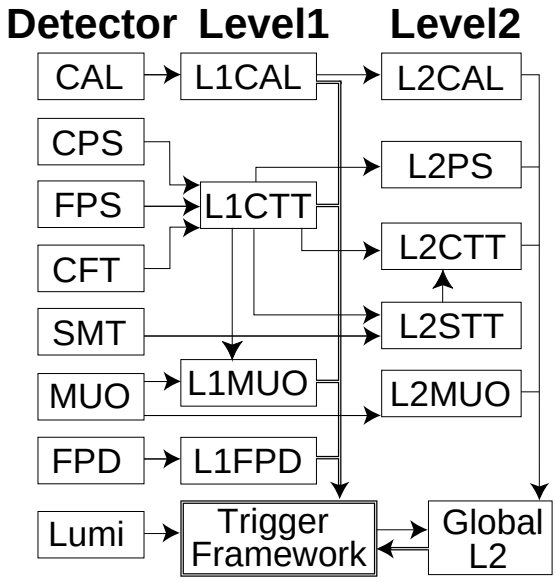


Figure 3.12: Overview of the DØ trigger and data acquisition systems [86].

Level 1 Trigger

The L1 trigger is implemented in specialized hardware, which makes the trigger decision within 3.5 μs. It consists of the Trigger Frame Work (TFW), which coordinates vetos, manages prescales of triggers and correlates triggers. The L1 calorimeter trigger (L1CAL) checks energy depositions exceeding the trigger levels in transverse energy deposits. The L1 central track trigger (L1CTT) reconstructs trajectories using the scintillator-based detectors. The L1 muon trigger (L1Muon) checks patterns for muons using hits from the muon wire chambers, the muon scintillator counters, tracks from L1CTT and the level 1 preshower detector (L1FPD).

Level 2 Trigger

The L2 trigger detector uses specific preprossessing engines. At the global stage (L2Global), correlations between subsystems are checked and input rates up to 10 kHz are computed. The maximum accept rate for L2Global is about 1 kHz. L2 can already combine data to form higher quality physics objects. Events selected by L2Global are based on 128 selections applied on L1 and script controlled criteria. Preprocessors include tracking, calorimeter, preshower and muon systems.

The L2 calorimeter trigger (L2CAL) identifies jets, electrons/photons, and calculates $\not\!\!E_{\mathrm{T}}$ using the data from the 2560 calorimeter towers. The jet algorithm

operates by clustering $n \times n$ (currently $n = 5$) groups which are centered on seed towers. Those with $\not\!\!E_{\mathrm{T}} \geq 2\,\mathrm{GeV}$ are stored with decreasing order in p_{T}. The electron/photon algorithm begins by a E_T ordered list of EM towers with E_T above 1 GeV. The largest neighboring towers are combined to build an EM cluster. L2 calorimeter $\not\!\!E_{\mathrm{T}}$ applies the vectorial summed E_T from individual trigger towers (from L1). Different minimum tower E_Ts and η ranges are possible.

The L2 muon trigger (L2Muon) uses calibration, a more precise timing and the L1Muon output. The central and forward preshower detectors (L2PS) provide high electron detection efficiency, electron photon separation, and high background (charged hadrons) rejection at the trigger level.

The L2 central track trigger L2CTT takes input from L1CTT and L2STT (L2 SMT trigger).

Level 3 Trigger

The level 3 trigger is a farm of Linux PCs, all of which reconstruct one event at a time. By using a faster algorithm than the offline reconstruction the whole event is analyzed and information like $\not\!\!E_{\mathrm{T}}$ or b-tagging probabilities are used. Correlations between different subdetectors can be used to make precise decisions if the event should be recorded. The output or accept rate of the farm ranges from 50-100 Hz depending on the $\mathcal{L}$ correlated with the time in store. Luminosity and typical rates for a store are given in figure 3.13. The numbers in the plot indicate the change of run transitions and prescales, as can be seen easily by the increased data taking rate after the transition.

Data Acquisition

The whole chain taking information and data from the first readout crate to the store tapes is called data acquisition. The storing of the data events which are finally accepted to tape is handled by several servers forming the level 3 data acquisition (L3DAQ) as seen in figure 3.14. The triggering and data acquisition is controlled by a coordination program called COOR [92].

As can be seen in figure 3.13, the rates increase from time to time during a store, while the luminosity drops exponentially. Triggers can be configured to accept only every n^{th} event, called prescale, where n is a natural number. Prescales are used to reduce the number of events which are less important, whereas unprescaled triggers are applied to gain as much information as possible from the delivered luminosity. For example, triggers for B physics have larger prescales at high luminosity, while at low luminosity the number of e.g. top candidate

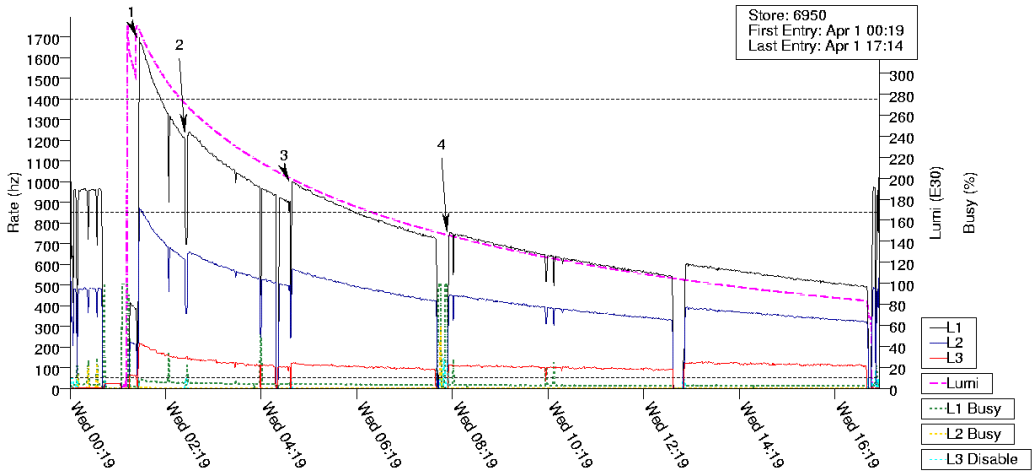


Figure 3.13: Luminosity and trigger rates. The numbers indicate the change of runs, called run transitions. The increased data taking event rate after the trigger set change is also visible. [91].

events is low enough to tolerate a higher number of b candidates to be recorded. Independent of the individual trigger prescales and current instantaneous luminosity, the complete trigger set is configured to have a final L3 accept rate of $\sim 50 - 100\,\text{Hz}$.

3.2.5 Luminosity Monitor

The luminosity monitor (LM) is installed within the DØ detector to provide the Tevatron luminosity at the DØ interaction region by measuring inelastic $p\overline{p}$ scattering. Additionally the LM measures the beam halo rates and the z coordinate of the interaction vertex.

The LM detector consists of two arrays of 24 plastic scintillators with photon multiplier tubes (PMT) readout located at $z = \pm 140\,\text{cm}$ as shown in figure 3.15. The detector is located in front of the EC and covers the pseudorapidity from $2.7 < |\eta| < 4.4$.

The luminosity $\mathcal{L}$ is measured by calculating the average number of inelastic collisions per beam crossing. To distinguish the beam halo background from $p\overline{p}$ collisions, the z-vertex position is calculated. The time of flight resolution is about 0.3 ns, where the dominant contribution is the variation in the light path length within the scintillation counters. The requirement for a beam beam collision is $|z_v| < 100\,\text{cm}$, which includes almost every Tevatron collision which has $\sigma_z \approx 30\,\text{cm}$. Beam halo particles have $z_v \approx \mp 140\,\text{cm}$, which are sorted out

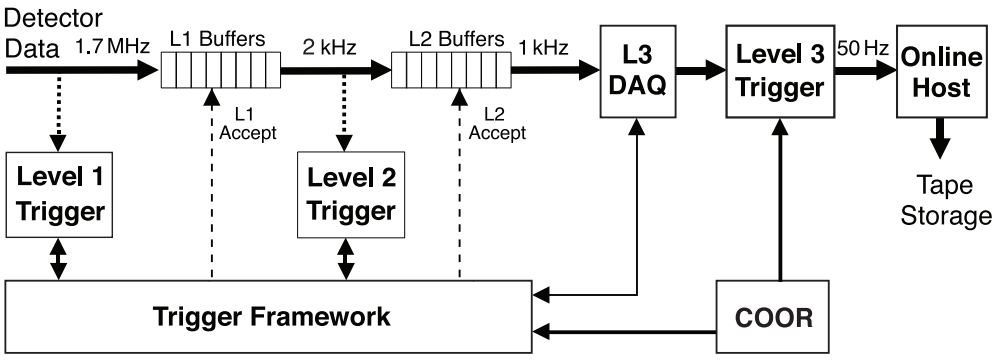


Figure 3.14: The DØ trigger and data acquisition systems [81].

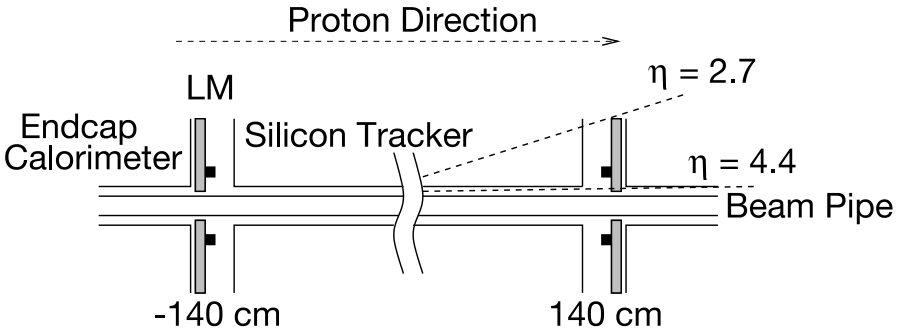


Figure 3.15: Location of the luminosity monitors on the z-axis [81].

by the $|z_v| < 100\,\text{cm}$ requirement.

The luminosity block is the fundamental unit in time for the luminosity measurement. The luminosity block number (LBN) is an index which monotonically increases throughout Run II. The LBN is incremented on run or store transitions, trigger framework or serial command link initialization, by request or after 60 seconds have elapsed. This time period is short enough to have an effectively constant instantaneous luminosity. Raw data is sliced into files corresponding to each LBN as is the calculation of the instantaneous luminosity.

CHAPTER 4

Object Identification and Reconstruction

After obtaining all information from the detector components, the signals have to be transformed to physical objects. These objects, their reconstruction and the respective selection criteria are described in the following chapter. The knowledge is needed in order to reconstruct fundamental objects like the primary vertex, electrons, muons, jets and missing transverse energy $\not{E}_{\mathrm{T}}$.

4.1 Reconstruction of Tracks and Primary Vertices

At first, the reconstruction of the hits charged particles leave in the inner detector are discussed and their assignment to the particle trajectories called tracks. Afterwards, tracks are combined, if they intersect at a certain point in space. These points are called vertices. Here, the so called primary vertices are discussed, which mark points where partons from the beam-beam collisions have interacted.

4.1.1 Tracks

When charged particles travel through the inner detector material, they loose energy through ionization and therefore deposit small amounts of energy along their trajectory. The track of a particle can be reconstructed by combining measured hits to a flight path or trajectory, using so called pattern recognition algorithms.

In the silicon microstrip tracker signals from silicon strips above a threshold level are combined to form a cluster. For the position of a particle the charge-

weighted mean of the hits of a cluster is taken to obtain a finer resolution than with the hits alone. The stereo angle of the silicon strips described in the last chapter allows the positioning of the location on a single strip. As the silicon detector resides in a 2 T magnetic field, the charge is not traveling straight to the read-out electrode. They drift with a Lorentz angle depending on the magnetic field strength and the Hall mobility of the electron or hole in the silicon. The measurement of the hit position is corrected for this drift angle.

In the central fiber tracker signals from fibers above the threshold in each doublet layer are also combined into clusters. Again, the stereo information is used to obtain the z coordinate of the hit.

Particles within a magnetic field travel on helical trajectories. This path needs to be corrected for field deviations and material in the detector, which are known quantities. DØ uses two algorithms to reconstruct tracks: the Histogram Track Finder (HTF) [93] and the Alternative Algorithm (AA) [94]. Both algorithms are run sequentially and return a number of track candidates as a result.

The HTF algorithm parametrizes the detector in ϕ and the track curvature in ρ.Hits in the detector for a curved track then fall into one bin in a $\phi-\rho$ histogram. Peaks in this histogram give track candidates. This is further enhanced by using a Hough transform [93], which maps each hit into a line in the parameter space. The intersection of lines can then be interpreted as a track. The HTF algorithm has a better performance on high p_{T} tracks.

The AA starts with three hits in the SMT. These lie along a trajectory originating from the interaction region. The trajectories are propagated outwards to the other SMT layers and later CFT layers. Given the new hit does not decrease the track quality, it is added to the trajectory. If there are multiple hits close to the trajectory, all are taken and form separate track candidates. Tracks with less than three hits in the SMT are reconstructed using the primary vertex candidates found with track candidates with at least three hits in the SMT. Track candidates with hits only in the CFT must originate near a reconstructed vertex.

Hits in the detector used by both algorithms are assigned to the longer track, i.e. the track with more hits, and are removed from the other track. In case both of the tracks have the same number of hits, the track with the smaller χ^2 keeps the hit. Afterwards tracks with too few hits and poor fitting (large χ^2) are removed and refitted using the Kalman smoothing [95] to gain the final tracks.

4.1.2 Primary Vertex

The reconstructed tracks are now used to find locations where the $p\bar{p}$ collisions have taken place, called primary vertices (PV). As there is more than one $p\bar{p}$

collision expected per beam crossing, usually multiple primary vertices are reconstructed. Vertices coming from decays of long lived particles like B hadrons are called secondary vertices (SV) and will be discussed in section 4.4.2 of this chapter.

DØ uses an adaptive vertex fitting algorithm [96] to reconstruct primary vertices. The first step is to use tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and at least two SMT hits in the SMT fiducial region. A z clustering algorithm combines all tracks within 2 cm of each other. The tracks in each z cluster are fitted to a common vertex by a Kalman filter algorithm [97], to determine the beam spot position. This algorithm then removes tracks with the highest χ^2, until the total vertex χ^2/ndf is smaller than ten.

Tracks in the z clusters are preselected using the distance of closest approach (dca) to the beam spot position calculated before. The distance is computed using the beam spot position and error, and is required to be smaller than five standard deviations.

The primary vertex is then calculated iteratively. All primary vertex candidates are fitted with the Kalman filter algorithm. Afterwards each track is weighted according to its χ^2 contribution to the vertex:

$$w_i = \frac{1}{1 + e^{\chi_i^2 - \chi^2_{\mathrm{cutoff}}/2T}} \,, \tag{4.1}$$

where χ_i^2 is the χ^2 of the track i to the current vertex, χ^2_{cutoff} is the χ^2 where the weight drops to 0.5, and the parameter T controls the sharpness of the function. $T > 0$ allows the track to contribute to both the PV and SV, though with a weight smaller than one. The parameter T is also called temperature in analogy to the Fermi function in statistical thermodynamics. Now the Kalman fit is recomputed as are the weights, until the latter are stable. Tracks with a weight smaller than 10^{-6} are eliminated from the procedure.

Finally, the hard scatter vertex has to be selected from the list of reconstructed PVs. In general the minimum bias interactions have a smaller transverse momentum than the hard scattering. For each track the probability whether it originates from a minimum bias vertex, is calculated using its $\log_{10} p_{\mathrm{T}}$. The individual probabilities are combined for all tracks of each PV. The vertex with the lowest probability of being a minimum bias vertex is chosen as the hard scatter primary vertex [98, 99]. The PV reconstruction efficiency was shown to be almost 100%.

4.2 Lepton Identification

The following sections will describe the identification of the leptons, needed for the semileptonic decay channel. The leptons are crucial for the trigger requirement and split the dataset in two parts: the electron+jets and the muon+jets channel. Taus are either coped by the decay into a lepton and treated as such or by decaying into hadrons and are not included in this analysis.

4.2.1 Identification of Electrons

Narrow clusters in the calorimeter are identified as EM objects or electron candidates. As narrow clusters count all towers inside a cone within a radius of $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$ around a seed tower. The p_{T} of the cluster is required to be larger than 1.5 GeV. Though all of the calorimeter systems are used to determine the energy of an electron, most of its deposited energy, E_{em} is required to be in the electromagnetic calorimeter thus, requiring the electromagnetic energy fraction f_{em} to be:

$$f_{\mathrm{em}} = \frac{E_{\mathrm{em}}}{E_{\mathrm{tot}}} > 0.9, \tag{4.2}$$

where E_{em} is the energy deposited in the electromagnetic part of the calorimeter and E_{tot} is the energy deposited in all calorimeter parts.

The electron candidates have to be isolated in $\eta \times \phi$ from other energy depositions. A second cone with a radius of $R = 0.4$ is defined. Up to 15 % of the total energy within the larger cone is allowed to be outside the smaller one, giving the following isolation criteria f_{iso}:

$$f_{\mathrm{iso}} = \frac{E_{\mathrm{tot}}(R < 0.4) - E_{\mathrm{em}}(R < 0.2)}{E_{\mathrm{tot}}(R < 0.4)} < 0.15. \tag{4.3}$$

The shower development of electrons differs from those of hadronic objects. Therefore, a 7×7 covariance matrix, the so called H-matrix, is calculated that quantifies how similar the EM object is compared to an electron using seven correlated variables. The variables are the deposited energy in the four EM layers, the total shower energy in the EM calorimeter, the z position of the primary vertex divided by its uncertainty, and width of the shower in $r-\phi$ in the third EM layer. The latter is chosen, since the shower maximum is located there and the EM calorimeter was built to have the finest granularity at this point as described in 3.2.2. The χ^2 calculated with the H-matrix should be small for electrons, with $\chi^2_{\mathrm{hmx7}} < 50$. It is required, that a track to the EM object in the

calorimeter is reconstructed and points to the EM object within $\Delta\eta < 0.5$ and $\Delta\phi < 0.5$, as an electron is a charged particle. The position is again determined in the third EM layer. For this a variable $\chi^2_{\mathrm{EM-trk}}$ is calculated:

$$\chi^2_{\mathrm{EM-trk}} = (\frac{\Delta\phi}{\sigma_\phi})^2 + (\frac{\Delta \mathrm{z}}{\sigma_\mathrm{z}})^2 + (\frac{E_\mathrm{T} - p_\mathrm{T}}{\sigma_{E_\mathrm{T}}})^2 \quad (4.4)$$

$$= \chi^2_{\mathrm{spatial}} + (\frac{E_\mathrm{T} - p_\mathrm{T}}{\sigma_{E_\mathrm{T}}})^2 \ . \quad (4.5)$$

This is a description of a so called "loose isolated" electron or loose electron. A "tight isolated" electron or tight electron needs to fulfill the loose isolation criteria and the output of the electron likelihood L_7 [100, 101] must be greater than 0.85. L_7 is a likelihood built from seven variables. The electron likelihood discriminates "true" electrons from fake electrons produced by background processes. The seven variables are: the electromagnetic energy fraction f_{em}, the χ^2 of the H-matrix, $E_\mathrm{T}/p_\mathrm{T}$, the probability $P(\chi^2_{\mathrm{spatial}})$, the z position of the *dca* of the matched track to the PV, the number of tracks around the matched track within a cone $\Delta R = 0.05$, and the sum of all track p_Ts in a cone of $0.05 < \Delta R < 0.4$, excluding the candidate track.

4.2.2 Identification of Muons

For the reconstruction of the muons, signals from the muon system and the central tracker are used [102, 103]. The muon system provides unambiguous information, about muons which passed the system. The tracking system has a more precise momentum resolution and a higher track finding efficiency.

In each muon layer the scintillator and wire hits are reconstructed to tracks. Muons found in this way are called local muons. Local muons that have a matched track from the tracker are called central track matched or global muons. The additional length of the track for global muons allows a much better measurement of the track curvature. Therefore most analyses use global muons. The calorimeter is used by looking at minimum ionizing particles (MIP).

The muons are classified into several categories depending on their number of hits in the muon system and the quality of the matched track [104]. The muons used for this analysis have to fulfill the muon quality "MediumNSeg3" and the track quality "medium" which translates into:

- $n_{\mathrm{seg}} = 3$, which means that the local muon must have hits in all three muon layers

- The "medium" quality requires at least two layer A wire hits, one A layer scintillator hit, two wire hits in the combined BC layer and one scintillator hit in the BC layer. In case of more than four BC wire hits, there is no requirement on the BC scintillator.

- The muon must be matched to a central track (global muon).

- The central track must be of medium quality, which means that $dca < 0.02$ with respect to the PV, if there is at least one SMT hit or $dca < 0.2$, if no SMT hits are found. The χ^2/ndf must be smaller than four.

- Cosmic muons are rejected by requiring muons to be inside a time window of ± 10 ns around the beam crossing.

An isolation cut is applied requiring the separation of the muon from hadronic objects. Muons are required to be isolated in $\eta \times \phi$ with $\Delta R > 0.5$ to any hadronic objects. Those muons are called loose isolated muons. Muons fulfilling the following criteria in addition to the loose ones are called tight isolated muons:

- $Rat11 = Halo(0.1, 0.4)/p_{\mathrm{T}_\mu} < 0.08$, where p_{T_μ} is the transverse momentum of the muon and $Halo(0.1, 0.4)$ is the sum of transverse energies in the calorimeter in a hollow cone with $0.1 < R < 0.4$ around the muon. For the calorimeter only electromagnetic and fine hadronic cells are considered. This is equivalent to 92 % of the energy, the muon deposits in the calorimeter, located within the smaller cone of 0.1.

- $Rattrk = TrkCone(0.5)/p_{\mathrm{T}_\mu} < 0.06$, where $TrkCone(0.5)$ is defined as the summed transverse momenta of all tracks within a cone of the radius $R = 0.5$ around the muon, excluding the muon itself.

4.3 Reconstruction of Jets

Hadronizing partons, quarks or gluons, form particles which tend to fly in the same direction as the original parton. In the detector, these particles form clusters of energies in the calorimeter and/or bundles of tracks in the tracking system. In the following, the term jets will refer to jets measured with the calorimeter also called calorimeter jets. In contrast, jets built by tracks will be called track jets. These are needed to identify heavy flavor objects. The algorithms to reconstruct jets, their quality, the correction of the energy between the original parton and the jet, and their resolution are described in the following sections.

4.3.1 Jet Algorithms

The jet algorithm is built to group particles for matching the kinematic properties of the jets as best as possible to those of the originating partons. The ideal jet algorithm should therefore have the following main attributes [105]. The algorithm should be *fully specified*, theoretically well behaving, meaning infrared and collinear safe, detector independent and order independent. From the theoretical point of view, the first three features are required, from the experimental one, the last four features are much more important:

- Infrared safety: The algorithm should not be sensitive to soft radiation in the event. Figure 4.1 shows an example for sensitivity to soft radiation, which leads to a different number of jets being reconstructed.

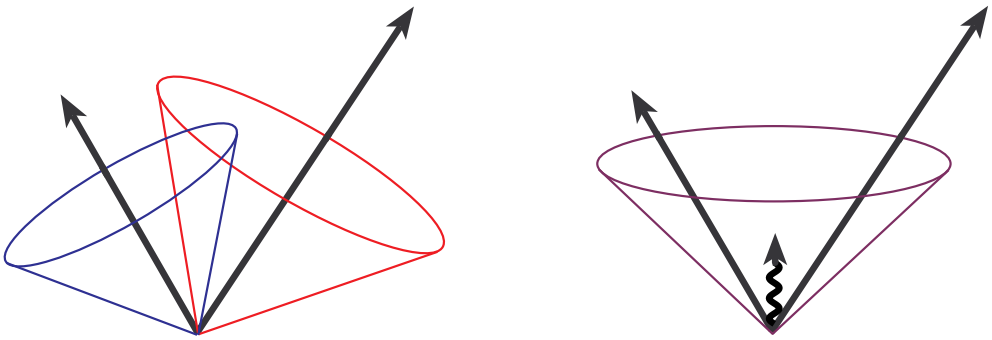
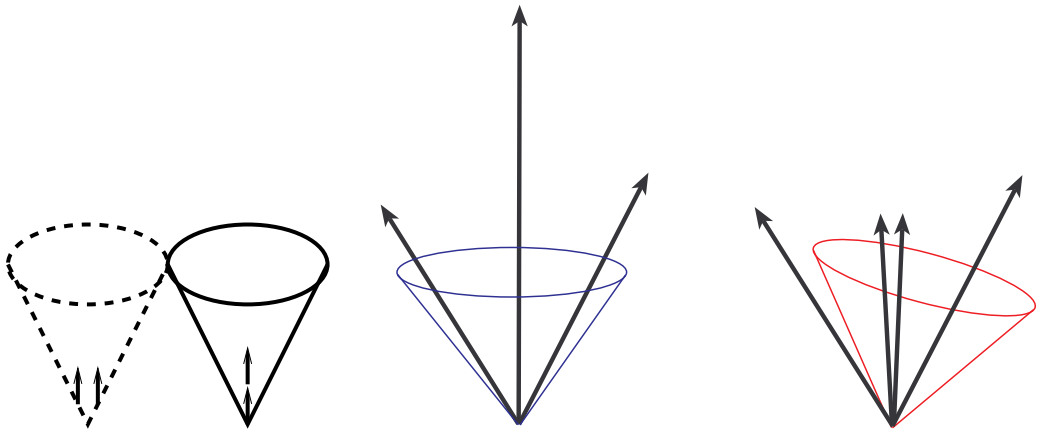


Figure 4.1: Example for soft radiation, which changes the jet reconstruction [105].

- Collinear safety: The algorithm should be robust against collinear radiation in the event. These effects can occur with jet algorithms that use seeds, where the minimal energy for a seed is below a particular threshold or with algorithms that are sensitive to the ordering of the particles, e.g. an ordering in E_{T}. Examples can be found in figure 4.2. In the left figure a jet having an energy above the threshold energy E_{cut} would not be reconstructed if it splits by collinear radiation into two particles both of which are below E_{cut}. The right side illustrates how the "rim particles" become more energetic than the central one, which is reduced due to collinear radiation. This leads to the use of the right particle before the middle one as starting point for the algorithm. The left particle is left out of the reconstruction, due to the limited cone radius.

  Jets in DØ data have been found to be 100 % efficiently reconstructed by using a seed tower threshold of 1 GeV [106]. The collinear dependence is removed when jets have a sufficiently large E_{T} to be reconstructed with 100 % efficiency.



(a) (b)

Figure 4.2: Figure (a) shows an example of sensitivity to ordering of particles in E_{T} while (b) is an example for collinear sensitivity for algorithms with seeds [105].

- Order independent: The algorithm should find the same jets at parton, particle, and detector level.
- Invariant under boosts (along the beam axis): The algorithm should give the same result independent of the boost of the event. This is especially important for hadron accelerators, which have boosted events in the laboratory frame in most of the cases. This is important for both the theoretical and experimental point of view.
- Detector independent: The dependence of the algorithms on the detector properties, like segmentation, response, or resolution should be as low as possible.
- Stability with luminosity: The reconstruction of the jets should be independent of the number of underlying inelastic interactions, which are increasing with luminosity.
- The algorithm should find all physically interesting jets, meaning the jets from high energy partons.

Besides those immanent features, the algorithm should also be easy to calibrate, easy to use, efficient in computer resources, and, of course, needs to be fully specified.

At the moment, two general ideas for reconstruction are available: The cone and the k_{T} algorithm, which are described in the following.

The Cone Algorithm

The basic idea of cone algorithms [107] is to group objects together which are inside a certain angular distance given by $\Delta R = \sqrt{\Delta y^2 + \Delta\phi^2}$. DØ uses the so called "Improved Legacy Cone Algorithm" (ILCA or Run II Cone Algorithm) [105, 108, 109].

Starting with objects, either partons, hadrons or calorimeter segments, objects within a certain radius R are merged and called protojets. Additional objects are added, until inside the cone no more objects can be found. The direction of the protojets will be recalculated after each recombination and the process iterated until a stable solution is found. All stable solutions will then become jets.

The most complicated part is to find a stable solution. The ideal solution would be the implementation of a sufficiently fine grid to generate protojets and combine those to stable jets, but this algorithm would use too much computing time. Therefore, the fact that jets possess a significant amount of their energy along their axis is used. So called seeds, which are calorimeter towers with a minimal threshold energy, are used. To reduce the number of seeds and save computing time, the DØ collaboration uses the ILCA algorithm. It combines multiple seeds to protojets until they reach an energy of 500 MeV. A complete overview can be found in [110].

The choice of seeds leads to problems with the radiation of soft particles, like two small angle soft particles seen in figure 4.2(a). To circumvent these, so called midpoints are introduced. They consist of the positions halving the distance between each combination of seeds. The midpoints are also used as starting points for the cones.

A remaining problem is the overlaying of multiple cones and the treatment of such cases. In this case the so called "splitting or merging" function is used. Whenever the percentage of transverse energy that the protojet with the smaller p_{T} shares with the other jet is bigger than a parameter f_{min}, 50 % within the DØ collaboration, the two jets are merged into a new jet (merging). Otherwise, the objects in the overlaying region will be assigned to the nearest jet (splitting). After this procedure, all jets with $p_{\mathrm{T}} > 8$ GeV will count as final jets.

The k_{T} Algorithm

Modern Monte Carlo generators use a hard scattering at the interaction point depending on the type of event. This is calculated with a QCD matrix element. The outgoing particles undergo the so called parton shower. Particles are radiated from the originating particle. Below a certain energy, particles hadronize into

colorless particles, which then might decay. This ends when all particles have been decayed to "stable" particles. This description is very successful and is equivalent to our understanding of the propagation of particles.

If the time line is inverted low energetic particles, which are near to each other, can be combined step by step. The k_{T} algorithm uses this ansatz. By doing so, collinear and infrared problems are solved, too. For theoreticians, this algorithm is a lot more attractive.

The algorithm starts with a list of so called preclusters, which can be taken from calorimeter cells, hadrons or partons. In an ideal world the algorithm could be used equally for partons, hadrons and calorimeter cells. At parton and hadron level, the starting four-vectors can be the partons and hadrons themselves. Living in the real world, particles might for example hit between two calorimeter cells. Therefore, the starting vectors are clustered with a granularity bigger than a single calorimeter cell, thereby eliminating the problem. Saving computing time is another reason as the k_{T} algorithm is rather CPU intensive. Recently, a slightly different version called Fastjet was developed by Salam which reduces the computing time drastically. For keeping the reconstruction time reasonable not all of the more than 45.000 cells could be used as starting four-vectors. They are therefore combined to towers before the preclustering takes place. Another reason to use preclustering is the real time pile off subtraction done in the DØ calorimeter. This makes negative entries possible for single cells, which are unphysical.

The k_{T} algorithm differentiates between several recombination schemes. The schemes regulate how the momentum of the new particle is calculated. The DØ collaboration decided to combine the particles by adding the four-vectors of the original particles. The starting particles are treated as massless. The cutting parameter at which the combination of particles stops, can be interpreted as a minimal angular distance in the momentum space, where $D = 0.4$ and $D = 1.0$ are used as cutting values for the DØ collaboration analyses.

The algorithm processes the following steps:

1. For each precluster i define: $d_i = {p_{\mathrm{T}_i}}^2$
   and for each pair of preclusters (i, j) with $i \neq j$ define:
   $$d_{ij} = \min(p_{T_i}^2, p_{T_j}^2)\frac{\Delta R_{ij}^2}{D^2} = \min(p_{T_i}^2, p_{T_j}^2)\frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{D^2} \ , \qquad (4.6)$$
   where $D \approx 1$ is a parameter of the jet algorithm. For $D = 1$ and $\Delta R_{ij} \ll 1$ d_{ij} is the minimal transversal momentum $k_\perp$ between the two preclusters.
2. Find the minimal d_i or d_{ij} and call it d_{min}.
3. If d_{min} is a d_{ij}, remove preclusters i and j from the list and add a precluster $(E_{ij}, \vec{p_{ij}})$ with:
   $$E_{ij} = E_i + E_j \qquad (4.7)$$

$$\vec{p}_{ij} = \vec{p}_i + \vec{p}_j. \tag{4.8}$$

This is called recombination.

4. If $d_{\min}$ is a d_i, remove precluster i and add it to the list of jets.
5. If there is at least one precluster left, go to step 1.

The algorithm produces a list of jets, having an angular distance of $\Delta R > D$. Figure 4.3 shows how preclusters are merged to jets.

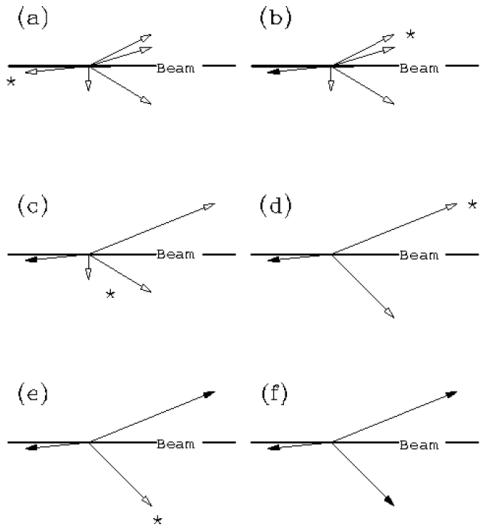


Figure 4.3: Simplified example of how jets are built with the k_{T} algorithm. Open arrows are preclusters, closed jets. Preclusters marked with an asterisk are the relevant ones for the next step[105].

The algorithm used by DØ differs from the one presented by Catani *et al.* [111, 112, 113]. The main differences can be found in the recombination scheme and the termination dependency. The recombination scheme in use for the applied algorithm within the DØ collaboration is the covariant E scheme (Equations 4.7 and 4.8), which is equal to adding the four vectors. It shows no energy defects and is the easiest one where a jet energy can be applied. The algorithm presented by Catani *et al.* introduces a termination parameter d_{cut}, which defines the minimal scale of the hard physics process. The event is split by this into a hard part and a part with small p_{T}(beam jets). The parameter can be chosen either ahead of the reconstruction, which is usually done, and if $d_{\min} > d_{\mathrm{cut}}$ the combining stops. All jets found so far having a $p_{\mathrm{T}}^2 < d_{\mathrm{cut}}$ will be declared as beam jets, while the rest will be used to reconstruct the hard scattering. Alternatively, d_{cut} can be set in each event so, that the number of reconstructed jets is the same in each event.

Ellis and Soper [114] present an algorithm which combines preclusters until they have a $\Delta R > D$. This is the choice of the DØ collaboration, as it gives a

similar separation of the jets as the cone algorithm. The use of d_{cut} would be well suited to define an exclusive jet algorithm.

The k_{T} algorithm is, however, not used in this analysis as there is no systematic study for the uncertainties on this algorithm done in the DØ collaboration. A complete study comparing cone and k_{T} jet algorithms at DØ can be found in [115]. Though the result favors the k_{T} algorithm, the lack of manpower available for the systematic study and the need for a development of a jet energy scale, makes it impossible to use it in this analysis.

4.3.2 Jet Identification

A set of quality cuts is applied to the jet candidates [116] so that the background from the calorimeter noise can be reduced and thereby distinguished from physical jets. This procedures also removes photon and electron candidates from the jet list. The jet cuts have been highly optimized in the recent analyses [116] and now depend on the jet location measured in the detector coordinates. The coordinates η_{D} and ϕ_{D} resemble the physical values of η and ϕ but are based on consecutive cell numbers. η_{D} ranges from –37 to +37 including zero and ϕ_{D} ranges from 0.5 to 64.5. An approximate conversion can be made by dividing the detector values by 10 to get the physical η values. The following cuts must be passed to qualify the candidate object as a jet:

- The electromagnetic fraction (f_{em}) is the amount of energy deposited in the electromagnetic part of the calorimeter divided by the total energy. Jets tend to have a uniformly distributed energy, leading to $f_{\mathrm{em}} \sim 0.5$. Very low values are unlikely, as particle jets have a significant amount of π^0's. Very high values are electrons and photons, which are misidentified as jets. Due to the detector geometry, especially in the ICD region, the cuts are optimized depending on η_{D}. A jet is required to have $f_{\mathrm{em}} < 0.95$ and:
  - $f_{\mathrm{em}} > 0.05$, or
  - $1.3 > ||\eta_{\mathrm{D}}| - 12.5| + \max(0, (40 \cdot \sigma_\eta - 4))$ (ICD region), where σ_η is the uncertainty on η , or
  - $f_{\mathrm{em}} > 0.03$ and $11.0 < |\eta_{\mathrm{D}}| < 14.0$ (ICD Region), or
  - $f_{\mathrm{em}} > 0.04$ and $|\eta| > 2.5$ (EC).
- The coarse hadronic fraction (f_{ch}) is the energy fraction of the jet deposited on the coarse hadronic part of the calorimeter divided by the total energy. Since jets do not deposit a large amount of energy in the last segment of the coarse hadronic calorimeter and the large size of the cells makes it prone to noise, they are not included in the total energy. The cuts are:
  - $f_{\mathrm{ch}} < 0.4$, or

- $f_{\mathrm{ch}} < 0.44$ and $|\eta| < 0.8$ (central region), or
- $f_{\mathrm{ch}} > 0.46$ and $1.5|\eta| < 2.5$ (EC excluding forward region), or
- $f_{\mathrm{ch}} > 0.6$ and $8.5 < |\eta_{\mathrm{det}}| < 12.5$ and $n_{90} < 20$,

where n_{90} is the number of calorimeter towers that represent 90 % of the total jet energy.

- The amount of energy measured by the first trigger step (L1) divided by the energy measured from the precision readout is the ($L1_{\mathrm{ratio}}$) defined as [117]:

$$L1_{\mathrm{ratio}} = \frac{p_{\mathrm{T}}(\text{from L1 readout})}{p_{\mathrm{T}}(\text{from precision readout})}. \tag{4.9}$$

The p_{T} from L1 is a scalar sum in a cone of the radius $R = 0.5$ from the 100 hottest L1 towers in the event. The p_{T} from the precision readout is a vector sum excluding the coarse hadronic layer [117].

- $L1_{\mathrm{ratio}} > 0.5$, or
- $L1_{\mathrm{ratio}} > 0.35$, and $p_{\mathrm{T}} < 15\,\mathrm{GeV}$, and $|\eta| > 1.4$ (EC), or
- $L1_{\mathrm{ratio}} > 0.1$, and $p_{\mathrm{T}} < 15\,\mathrm{GeV}$, and $|\eta| > 3.0$ (forward region), or
- $L1_{\mathrm{ratio}} > 0.2$, and $p_{\mathrm{T}} \geq 15\,\mathrm{GeV}$, and $|\eta| > 3.0$ (forward region).

4.3.3 Jet Energy Corrections

The sum of the measured energy of the objects finally qualifying as jets is not the energy of the particle jet. Several factors can change the measured energy of the jet. Absorbing material in front of the calorimeter, like the tracking system, magnets, cryostats, and the preshower, reduce the energy. Gaps in the calorimeter, like the ICD region or instrumentally not filled parts due to readout electronics or support structure, change the measured energy. The jet algorithm does not assign all energy deposits originating from the same parton to the jet, since some of these deposits fall out of the jet cone, which reduces the energy measured from a jet. On the other side the underlying event or multiple interactions add energy to the jet. The jet energy scale (JES) tries to parametrize these effects to correct to the original energy of the particle jets [118, 119]. The formula can be written as:

$$E_{\mathrm{jet}}^{\mathrm{ptcl}} = \frac{E_{\mathrm{jet}}^{\mathrm{raw}} - O}{F_\eta \times R \times S}\ , \tag{4.10}$$

where $E_{\mathrm{jet}}^{\mathrm{ptcl}}$ is the corrected jet energy, $E_{\mathrm{jet}}^{\mathrm{raw}}$ is the uncorrected jet energy, O is the energy correction offset, F_η is the relative response correction, R is the absolute response correction, and S is the showering correction.

The correction factors are derived separately as they differ in data and Monte Carlo. The order of derivation is given below, including the exact meaning.

- The offset energy correction O is subtracted from the raw jet energy. The contributions for this energy are jets, that do neither stem from the hard interaction process, nor from electronic noise, uranium noise, energy from previous bunch crossings (pile up), beam remnants, or additional minimum bias interactions of the same bunch crossing. The offset energy is measured in minimum bias interactions by summing up all calorimeter towers with the cone jet radius R and defining the energy inside the cone as the offset energy O. Figure 4.4 shows the dependence on the number of PVs, which results in a linear rising function of O, as expected. On the other hand figure 4.5 shows the offset energy as a function of η_{det} for different numbers of primary vertices.
- The relative response correction F_η is needed to correct the non-uniformity of the calorimeter especially between the CC and the EC, where the ICD region and massless gaps reside. After this correction the calorimeter should give a uniform response to energy deposition. The correction factor is determined by using the missing transverse energy projection method (MPF) in events with photons recoiling against exactly one jet [120]. The missing transverse energy and the p_{T} of the photon and jet are measured. The photon serves as a tag, while the jet probes the energy measured by the photon, illustrated in figure 4.6. The photon energy scale is measured very precisely and independently from $Z \rightarrow ee$. The jet response relative to $\eta_{\text{det}} = 0$ can then be measured from the E_{T} imbalance as a function of η_{det}. The correction factor F_η can be seen in figure 4.7, where the drop in the response can clearly be seen as dips in the distribution.
- The absolute response correction R corrects different responses of the calorimeter for hadrons and leptons, as well as the non-linearity of the calorimeter response to the particle energy. It is determined from γ + jet events like F_η. Figure 4.8 shows the absolute response correction as a function of the jet energy.
- The showering correction S corrects out of cone showering. This happens not only due to soft gluon emission but also due to the magnetic field inside the detector and the showering inside the calorimeter itself. It is measured by taking γ + jet events with exactly one photon, one jet, and one primary vertex. The deposited energies inside and outside the cone are measured and their ratio is used to calculate the correction. In simulated events, the ratio between the jet at the particle level and at the reconstruction level gives the correction. Fig. 4.9 shows the showering correction for data.

The uncertainty of the jet energy scale is one of the dominating systematic

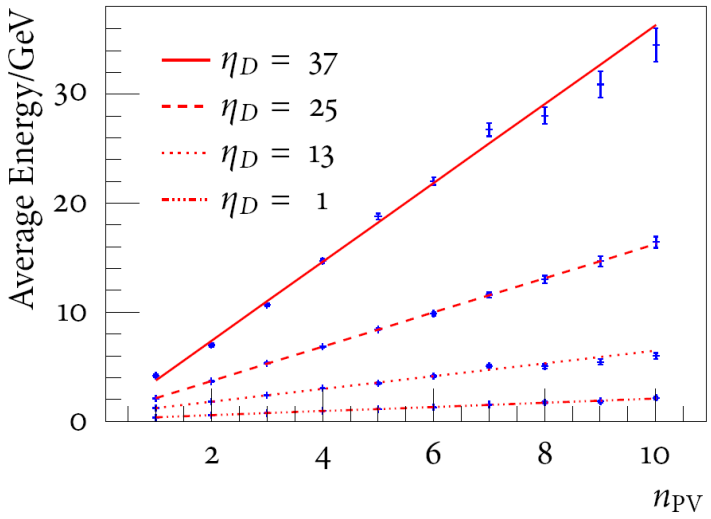


Figure 4.4: Offset energy for different PVs. The dependence for different PVs, which rises linearly with the number of minimum bias interactions is shown [118].

uncertainties for the $t\bar{t}$ cross section measurement in the l+jets channel. The relative uncertainty of the jet energy scale is shown for the preliminary JES in figure 4.10(a) and the final JES in figure (b). A huge effort was made to minimize the JES uncertainty. It was possible to reduce the relative uncertainty by more than half to a maximum of less than 4 % at its highest value. The major uncertainty still comes from jets within the ICD region and for low energy jets. For Run IIb the JES is still preliminary. Finally, jets with muons receive an additional correction for the energy loss from the muon and the neutrino.

4.3.4 Jet Energy Resolution

The jet energy resolution (JER) is based on data events [88]. The parametrization of the JER is given in equation 3.6 in section 3.2.2. Additionally a method called jet shifting, smearing, and removal (JSSR) is introduced [121, 122] to incorporate differences between Monte Carlo and data for jet resolution, jet reconstruction efficiencies, as well as identification efficiencies. After the JSSR recalibration, smearing and discarding Monte Carlo jets behave similar to jets from observed data.

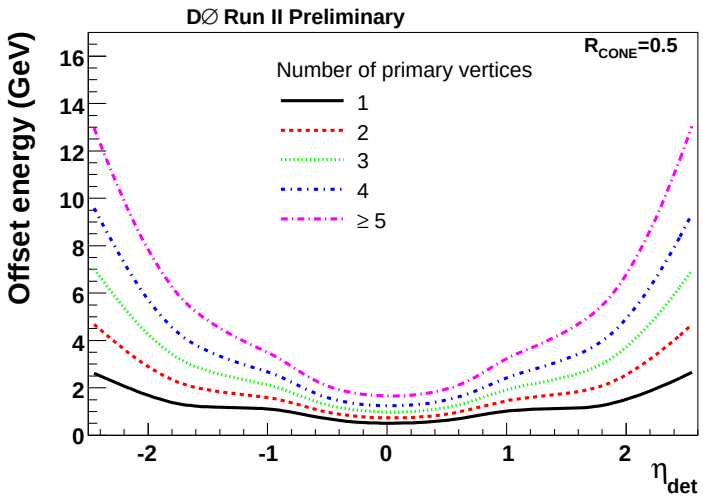


Figure 4.5: Offset energy for different η_{det}. The dependence for changing η_{det} is shown [118].

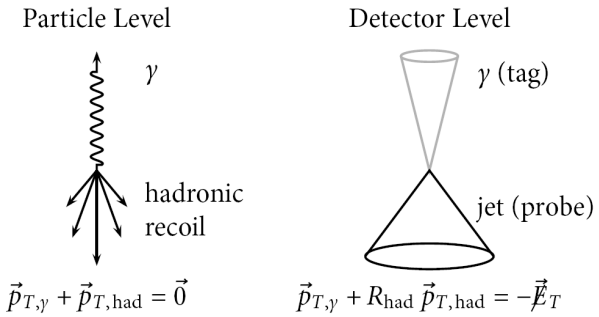


Figure 4.6: Illustration of the $\not{E}_{\text{T}}$ projection method in back to back γ + jet events [120].

4.4 Secondary Vertices & b-tagging

At the Tevatron B hadrons usually travel several millimeters before decaying. They thus form another vertex, from which the decay products emerge. Such a vertex is displaced compared to the PV discussed before. The observation of such a vertex called secondary vertex (SV) allows the detection of long living objects like b or c quarks.

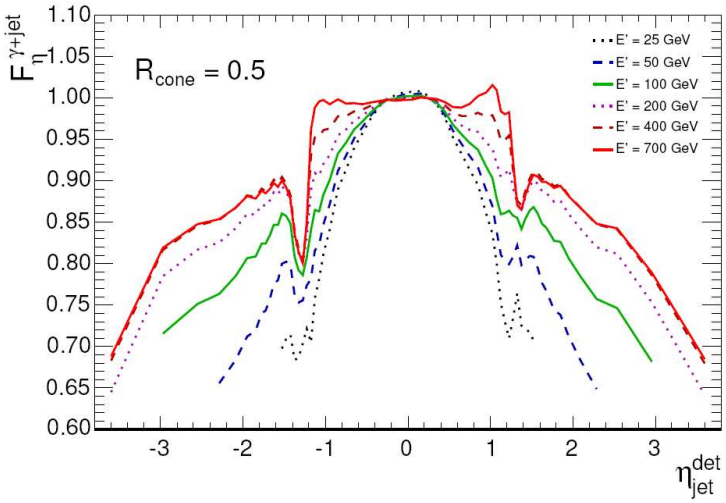


Figure 4.7: Relative response correction for JES in data [118]. The correction factor F_η is shown depending on the jet η. A low value means a low response from the calorimeter, as the factor is in the denominator of the final correction. The low response in the ICD reagion is clearly visible.

4.4.1 Secondary Vertex

The reconstruction of the secondary vertex is based on the Kalman filter algorithm [123].The first step is the reconstruction of track jets [124], which is independent of calorimeter input. The clustering algorithm starts with the construction of z-preclusters [125] like for the PV. The p_T ordered tracks are added to the z-precluster, if the dca in the z-coordinate of the track with respect to the precluster is smaller than 2 cm. At the preclustering stage, tracks are selected for each precluster. Each track needs at least one hit in the SMT, $p_\mathrm{T} > 5$ GeV, and the $|dca| < 0.2$ cm for the xy-plane and $|dca| < 0.4$ cm for z. Jets are formed from tracks with a simple cone algorithm with $R = 0.5$ and a seed track with $p_\mathrm{T} > 1$ GeV to reduce track jets from the underlying events.

The vertex reconstruction algorithm is applied to track jets where at least two tracks have a large impact parameter significance $dca/\sigma_{dca} > 3$ [126], where σ_{dca} is the uncertainty of the dca. The track criteria can vary depending on the tagger in use. This will be discussed in the next section. The algorithm fits all combinations of two tracks in each track jet. Additional tracks are attached to the vertex, if the χ^2 is below a tunable threshold. Once the threshold is reached, no more tracks are added to the vertex. Each track can be part of several vertices.

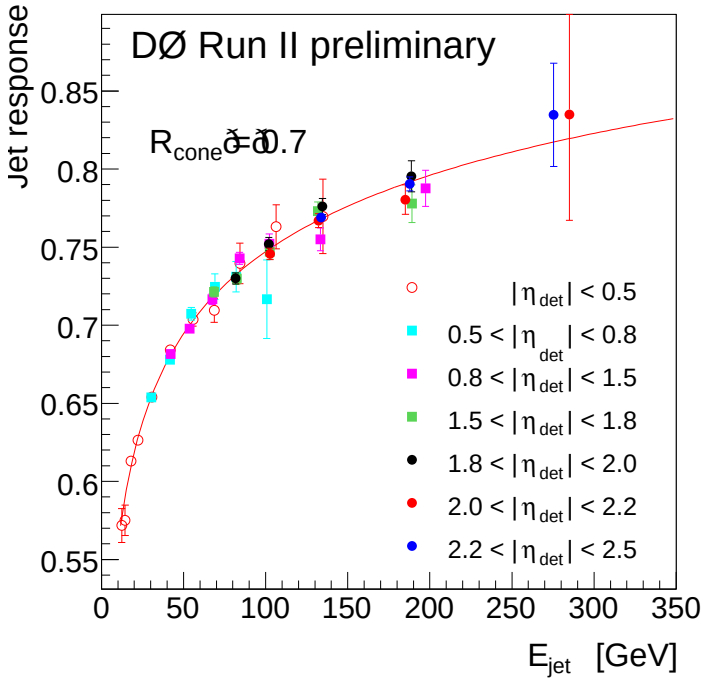


Figure 4.8: Absolute response correction for $R = 0.7$ [118].

4.4.2 Identification of b Jets

Each top quark decays with about 100 % probability into a W boson and a b quark resulting in $t\bar{t}$ events having at least two b quarks. This means that b quark jet identification with high efficiency is a major point in reducing the background for this analysis. Two different ideas are used at DØ to identify b quark jets. The first is that the b quark decays into a virtual W boson and a lighter quark, where the W boson then decays in about 20 % of all cases into a lepton, which can be identified. The other is that B hadrons travel a macroscopic distance of several millimeters before they decay, forming a secondary or displaced vertex.

DØ uses a so called soft lepton tagger, using the decay into a muon and three so called lifetime taggers exploiting the distance traveled by the b quark hadronization object, as well as a combination of those four in a neural network. The latter will be used in this analysis.

Taggability

Before it is possible to apply a b-tagging algorithm the calorimeter jet needs to be taggable [124]. This step is introduced to remove detector effects from the

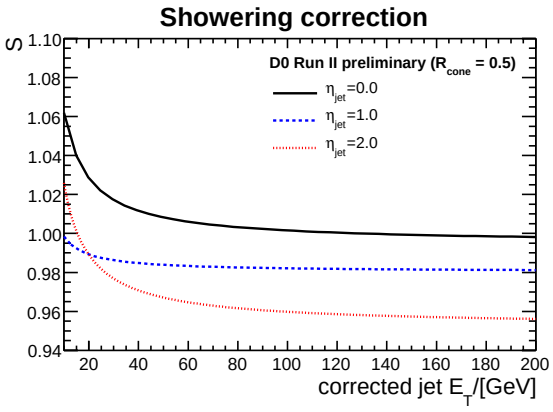
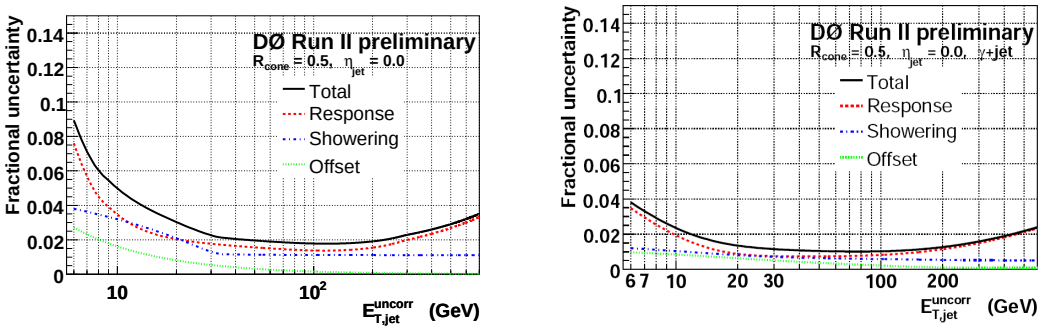


Figure 4.9: Showering correction versus corrected transverse jet energy [118].



(a) (b)

Figure 4.10: Relative uncertainty of preliminary (a) and final (b) Run IIa JES correction [118, 119].

b-tagging algorithms. This allows to parametrize these effects and use them in Monte Carlo simulations. A jet is taggable, if it matches a track jet within a cone $R = 0.5$ with the following requirements:

- at least one seed track with $p_{\mathrm{T}} > 1.0$ GeV,
- at least one other track with $p_{\mathrm{T}} > 0.5$ GeV,
- all tracks have at least one SMT hit,
- $dca < 2\,\mathrm{mm}$ in the xy-plane and $dca < 4\,\mathrm{mm}$ along the z-axis,
- Δz between the track precluster and the z of the position of the closest approach is $< 20\,\mathrm{mm}$.

The jets also need to be originating from the PV. The PV has to be reconstructed with at least four tracks, otherwise no jet in the event will be taggable. Since at least one registered hit is required to be from the SMT, jets have a strong dependence on the chance to be taggable, depending on the z-position of the PV (PV_z) and η. Figure 4.11 shows a visualization of this, including the sign of the

$\eta \times PV_z$ value. The latter is negative, if the jet is oriented towards the center of the detector and positive otherwise.

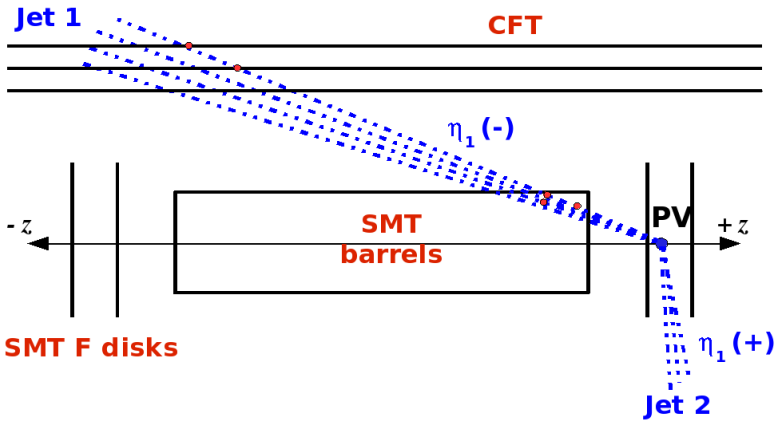


Figure 4.11: Region dependence of Taggability [127]. The probability to be taggable is high for jet 1 and low for jet 2 as the particles within the jet do not traverse a single pixel layer. Possible hits of the jet particles are marked with red dots.

Jets traveling back to the center, have a higher chance of passing silicon layers and therefore of being taggable. This can clearly be seen in figure 4.12 displaying the Taggability Rate Functions (TaggabilityRF). Those functions are needed, as the detector simulation in Monte Carlo does not yield the correct ratio for taggable jets. Therefore, the TaggabilityRFs are measured in data and applied to the Monte Carlo. The functions are parametrized in η and p_T of the jet, and additionally split into six regions with increasing z distance to the center of the detector. It can be seen, that jets having a high $|z|$ and traveling back into the detector, thus belonging to the minus region, for example the lower right plot in figure 4.12, have a higher chance to be taggable once they point inward. If one compares the electron parametrization to the muon one in figure 4.13, the dips in the ICD region show up in the electron parametrization. The muon distribution is flat, for example the upper left plot, which shows the outward pointing jets from the center of the detector, where η is similar to η_{det}.

The TaggabilityRFs also serve as a step to make the b-tagging algorithm sample and detector independent, as both of the dependences should be parametrized in the TaggabilityRF. This allows the b-tagging algorithms to be evaluated and compared mostly independently from any detector effects.

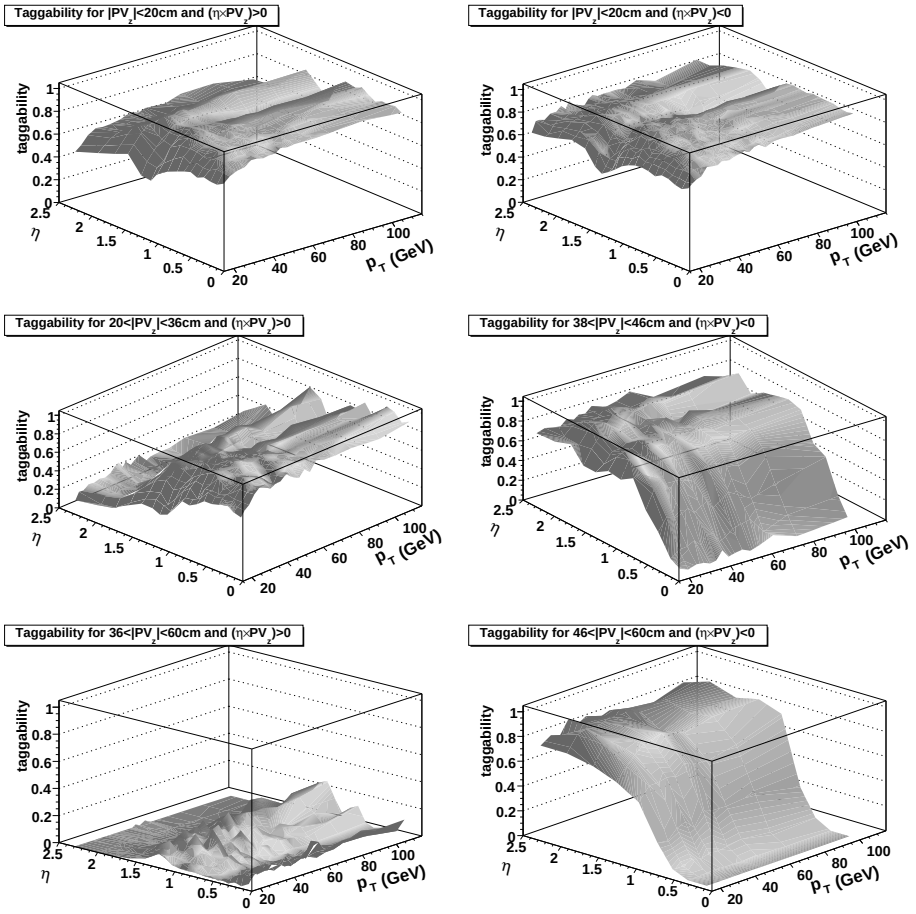


Figure 4.12: TRFs for the e+jets data set. The parametrization is done separately in six regions of $|PV_z|$ and $\eta \times PV_z$ [128].

b-Tagging

In the following paragraphs, the five different DØ taggers will be briefly explained, as well as their combination to a neural network tagger. Finally the use of the Tag Rate Functions (TRF), which are similar to the TaggabilityRFs presented in the last section is presented.

Soft Lepton Tagger In almost 20 % of all cases b quarks decay into a muon, either from the W boson from the b quark decay itself or from a decaying c quark, that was produced in the decay of the b quark. The soft lepton tagger (SLT) [129] identifies those muons. The advantage is that there is good agreement between Monte Carlo simulation and observed data. However, the b-tagging efficiency is low, due the fact that only 20 % of the b quarks have a muon in their decay chain.

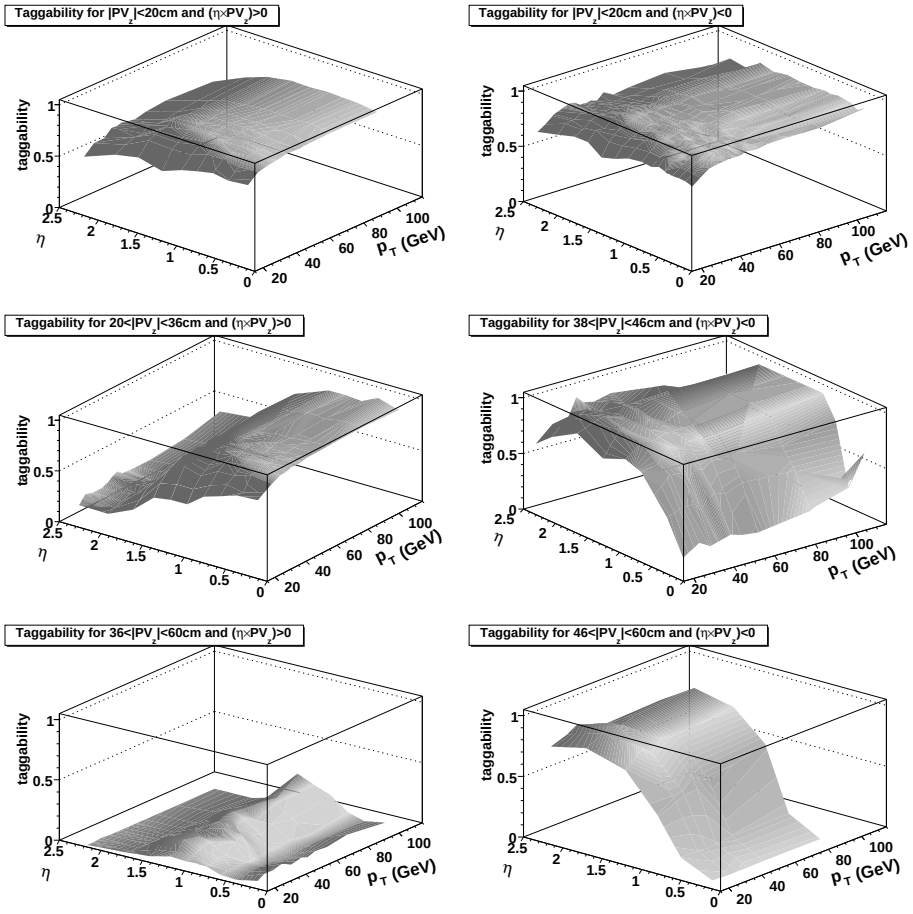


Figure 4.13: TRFs for the μ+jets data set. The parametrization is done separately in six regions of $|PV_z|$ and $\eta \times PV_z$ [128].

The SLT cannot distinguish between c and b quarks.

Counting Signed Impact Parameter Tagger The counting signed impact parameter tagger (CSIP) [130, 131] exploits the long lifetime of b quarks. Tracks from b quarks are usually displaced from the PV and have an impact parameter larger than zero as figure 4.14 shows. The sign of the projection of the *dca* is either positive, if the jet and the SV are on the same side or negative, if they are on opposite sides. Tracks from b jets have a larger positive impact parameter significance than light jets, as can be seen in figure 4.15. Jets are b-tagged, if more than three tracks with impact parameter significance of larger two or at least two tracks with impact parameter significance greater than three are found. Different working points can be defined by cutting the minimal p_{T} of the tracks. Increasing p_{T} decreases the fake rate and the efficiency.

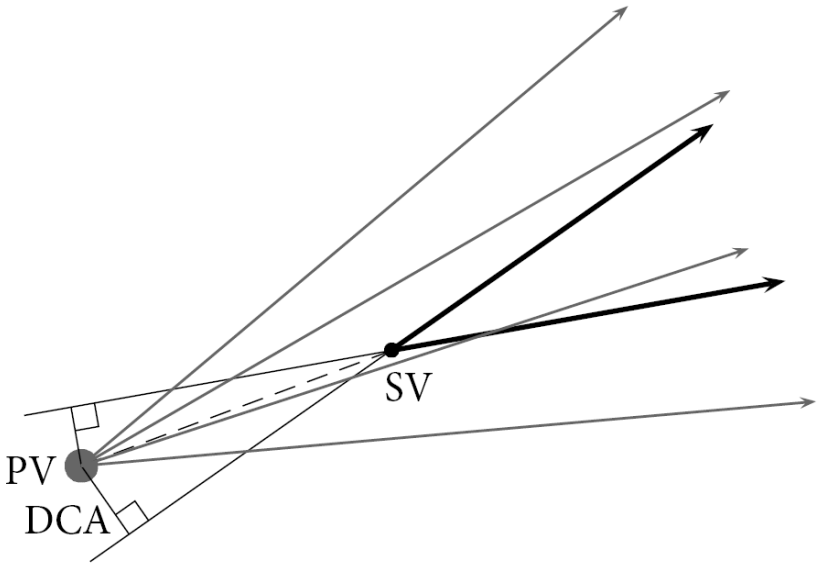


Figure 4.14: Illustration of the distance of the closest approach (*dca*) to the PV, also called impact parameter. The distance between PV and SV is plotted as a dashed line [131].

Jet Lifetime Probability Tagger The jet lifetime probability tagger (JLIP) [132, 133] uses the impact parameter information from all tracks to calculate the jet lifetime probability P_{jet}. P_{jet} determines the possibility that all tracks originate from the PV. This is flat for light jets, while b and c jets peak at low values. Cutting on the probability selects the heavy flavor jets. Different working points can be defined by cutting on various P_{jet}.

Secondary Vertex Tagger The secondary vertex tagger (SVT) [134, 126, 135, 136, 137] reconstructs the SV as described in section 4.4.1 and measures the corresponding decay length.

The secondary vertex is required to have:

- a $\chi^2/ndf < 15$ for each track,
- at least two SMT hits and $|dca| < 0.15\,\text{cm}$ in the xy-plane and $|z_{\text{dca}}| < 0.4\,\text{cm}$ for each track,
- a vertex $\chi^2/ndf < 100$,
- a maximum decay length in the xy plane ($L_{\text{xy}} < 2.6\,\text{cm}$) to remove vertices from long living particles like Λs or K^0_Ss,
- collinearity > 0.9, where collinearity is the product of L_{xy} and the sum of all track momenta normalized to one.

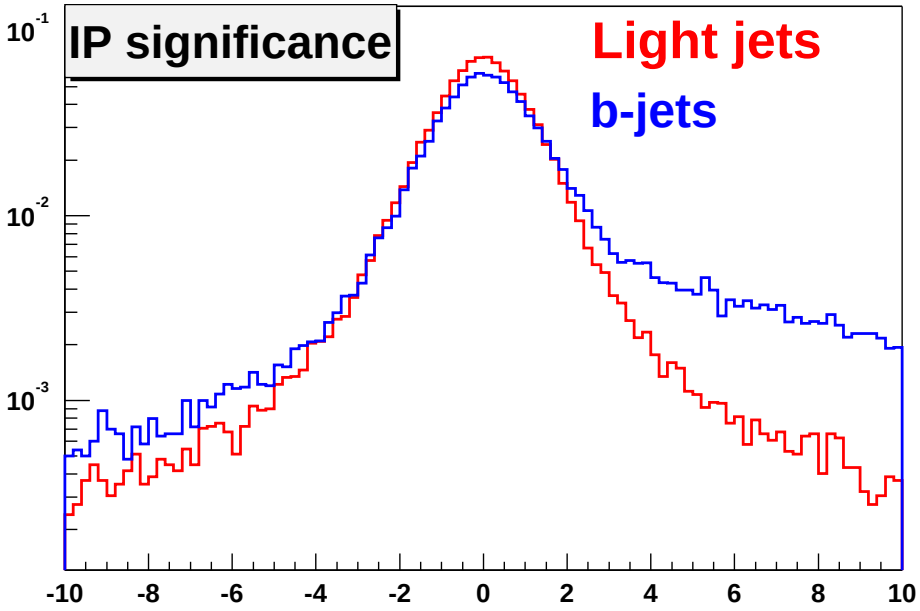


Figure 4.15: Illustration of the difference in the counting signed impact parameter significance for light and heavy flavor jets [131].

Jets are b-tagged, depending on the working point of the tagger given in table 4.1. b-tags with decay length L_{xy} smaller than zero are called negative, while those with $L_{\mathrm{xy}} > 0$ are called positive. The negative tagged jets originate from mis-reconstructed jets and can therefore be used to measure the fake tag rate. The tagger is optimized for low mis-tag rates which are 1%, 0.5% and 0.25% for the loose, medium and tight working point. The superloose working point is used as input for the neural network tagger, which will be discussed in the next paragraph.

Working Point	Track p_{T}	Track $dca/\sigma_{\mathrm{dca}}$	Track χ^2/ndf	Vertex $L_{\mathrm{xy}}/\sigma_{\lvert L_{\mathrm{xy}}\rvert}$
Tight	> 1.0 GeV	≥ 3.5	≤ 3.0	≥ 7.0
Medium	> 1.0 GeV	≥ 3.5	≤ 10.0	≥ 6.0
Loose	> 1.0 GeV	≥ 3.0	≤ 10.0	≥ 5.0
Loose_xtrack	> 0.5 GeV	≥ 3.0	≤ 10.0	≥ 5.0
Superloose	> 0.5 GeV	≥ 0.0	≤ 15.0	≥ 0.0

Table 4.1: Tag working points of the secondary vertex tagger [135].

Neural Network Tagger The Neural Network Tagger (NN tagger) [138, 139, 140] which is used in this analysis combines the JLIP, CSIP, and SVT into one neural network. The optimization done for the NN tagger gives the following variables as best choice to discriminate b jets from light jets, given in the order of their

separation power. The input variables are shown in figure 4.16 for observed and simulated QCD multijet events as well as for $b\bar{b}$ Monte Carlo events. The several variables are defined as following:

SVT_{SL} S_{dl}: is the decay length significance $S_{dl} = L_{xy}/\sigma_{L_{|xy|}}$ of the SVT for the superloose working point. In case that several SVs are found within the jet cone, the highest dls is used.

CSIP comb is a weighted combination of the impact parameter significance S_{ip} of the tracks. The combination consists of the sum of tracks with $S_{ip} > 3$, $S_{ip} > 2$, $S_{ip} < 3$, and $S_{ip} < 2$.

JLIP prob is the probability P_{jet} that the jet originates from the PV.

SVT_{SL} χ^2/ndf is the χ^2/ndf of the SV for the superloose working point.

SVT_{L} n_{tracks} is the number of tracks associated to the SV for the loose working point.

SVT_{SL} mass is the invariant mass of the SV for the superloose working point.

SVT_{SL} num is the number of SVs found within the jet cone using the superloose working point.

Different working points can be defined by cutting on the neural network variable. Figure 4.17(a) shows the output of the NN tagger for simulated $b\bar{b}$ and QCD multijet events. The separation power between the $b\bar{b}$ sample, which peaks at one, and the QCD sample, which consists mainly of light jets, can easily be seen. Table 4.2 lists the certified working points for the NN tagger. Though in principle all values can be used, only a couple of them have been approved, meaning certified for use by the DØ collaboration. This analysis uses the medium tag working point, which gives the best performance for semileptonic $t\bar{t}$ cross section analyses [141]. The working point has an efficiency of about 50 % while having a fake rate of less than 1 % which can be seen in figure 4.17(b). It also shows the comparison to the JLIP tagger.

Tag Rate Functions Since the Monte Carlo simulation again is not perfect in the simulation of tracking, the efficiencies for b-tagging do not match those in observed data. The same method as for taggability is used. The parameterizations used are called tag rate functions (TRF) and are parametrized in E_T and η. A separate TRF is made for the b-, c-, and mis-tag rates. All jets have to be taggable before TRFs can be applied.

The TRFs for b-jets are measured in muonic b-jets from the observed data, with a method called System8 [139, 142]. This method uses two different taggers and their detected b-tags in two samples. The second tagger used to determine the NN b-tag probabilities is the SLT tagger. This method yields the TRF for the

Operating Point Name	NN Output Cut
L6	> 0.1
L5	> 0.15
L4	> 0.2
L3	> 0.25
L2	> 0.325
Loose	> 0.45
oldLoose	> 0.5
Medium	> 0.65
Tight	> 0.775
VeryTight	> 0.85
UltraTight	> 0.9
MegaTight	> 0.925

Table 4.2: Certified tag working points of the neural network tagger [139].

muonic b-jets which are also derived in Monte Carlo. The scale factor $SF_{\mathrm{b}}(E_{\mathrm{T}}, \eta)$ is the ratio for muonic b-tag TRFs in observed data and Monte Carlo simulation. As the efficiency for inclusive b-tagging is not expected to be the same as for muonic b-tagging, the correction needed is believed to be a scale factor taken from Monte Carlo simulation:

$$\epsilon^{\mathrm{data}}_{\mathrm{incl\ b}}(E_{\mathrm{T}}, \eta) = \epsilon^{\mathrm{MC}}_{\mathrm{incl\ b}}(E_{\mathrm{T}}, \eta) \cdot SF_{\mathrm{b}}(E_{\mathrm{T}}, \eta) \ , \tag{4.11}$$

where $\epsilon^{\mathrm{data}}_{\mathrm{incl\ b}}(E_{\mathrm{T}}, \eta)$ is b-TRF for recorded data and $\epsilon^{\mathrm{MC}}_{\mathrm{incl\ b}}(E_{\mathrm{T}}, \eta)$ the b-TRF for simulated events.

For the c-TRFs the assumption is made that the difference between c and b quarks is the same in Monte Carlo simulation as in observed data. This allows to set $SF_{\mathrm{c}}(E_{\mathrm{T}}, \eta) = SF_{\mathrm{b}}(E_{\mathrm{T}}, \eta)$. This results in a similar equation as for the b-TRFs:

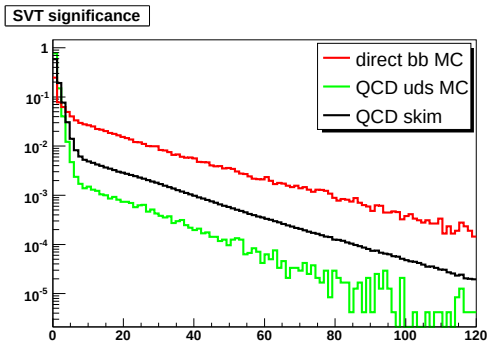
$$\epsilon^{\mathrm{data}}_{\mathrm{incl\ c}}(E_{\mathrm{T}}, \eta) = \epsilon^{\mathrm{MC}}_{\mathrm{incl\ c}}(E_{\mathrm{T}}, \eta) \cdot SF_{b}(E_{\mathrm{T}}, \eta). \tag{4.12}$$

The fake rate, which is measured by using the negative tag $\epsilon_{\mathrm{NT}}(E_{\mathrm{T}}, \eta)$ from the NN tagger, is supposed to give the fake or mis-tag rate in the positive tags. As there also are negative true b- and c-quarks in the negative tagged jets and light jets may also have long lived particles occasionally, the negative tag rate is corrected by two factors. The SF_{hf} is the ratio of the light jet to the total negative tag rate needed for the correction of the heavy flavor contribution. The ratio SF_{light} is the ratio of the light jet positive tag rate to the light jet negative tag rate. It is used to compensate for the amount of long living light particles,

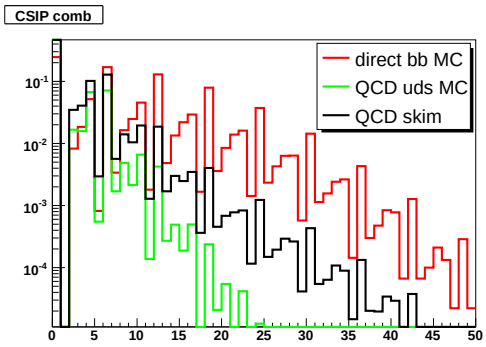
mainly K^0. As one needs to know the true quark content, both factors are derived from Monte Carlo simulations. This leads to the final light-TRF fake rate of:

$$\epsilon_{\text{light}}(E_{\text{T}}, \eta) = \epsilon_{\text{NT}}(E_T, \eta) \cdot SF_{\text{hf}}(E_{\text{T}}, \eta) \cdot SF_{\text{light}}(E_{\text{T}}, \eta) \tag{4.13}$$

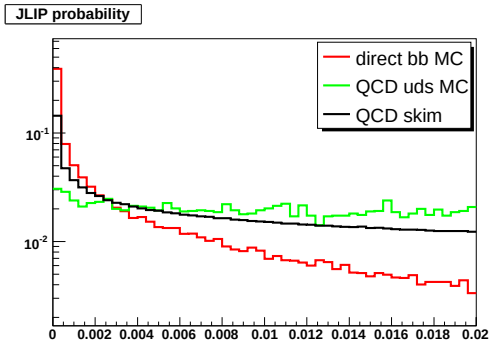
Figure 4.18 shows the projections for the b- and c-tag rate in E_{T} and η. The mis-tag rate is parametrized in E_{T} and calorimeter region (CC, ICD EC).



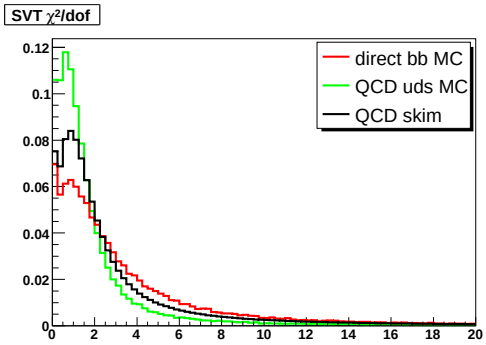
(a) compares the S_{dl} of the SVT,



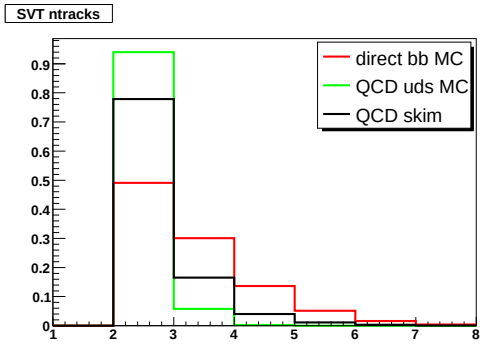
(b) the S_{ip} of the tracks,



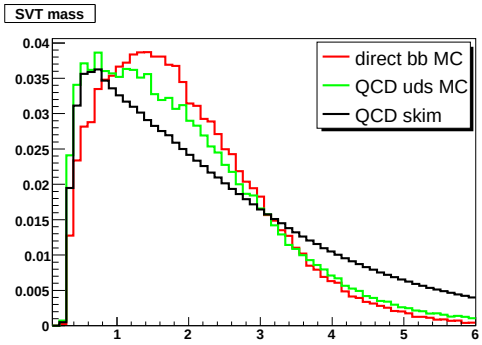
(c) the P_{jet} that a jet originates from the PV,



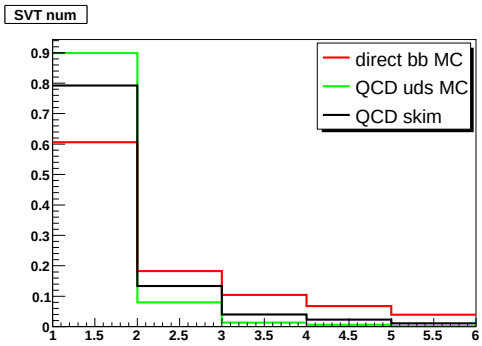
(d) the χ^2/ndf for the SVT,



(e) the number of tracks associated to the SV,

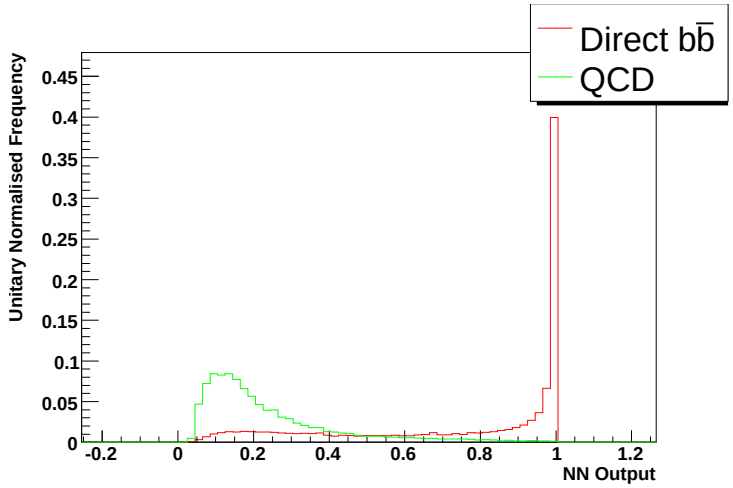


(f) the invariant mass of the SV,

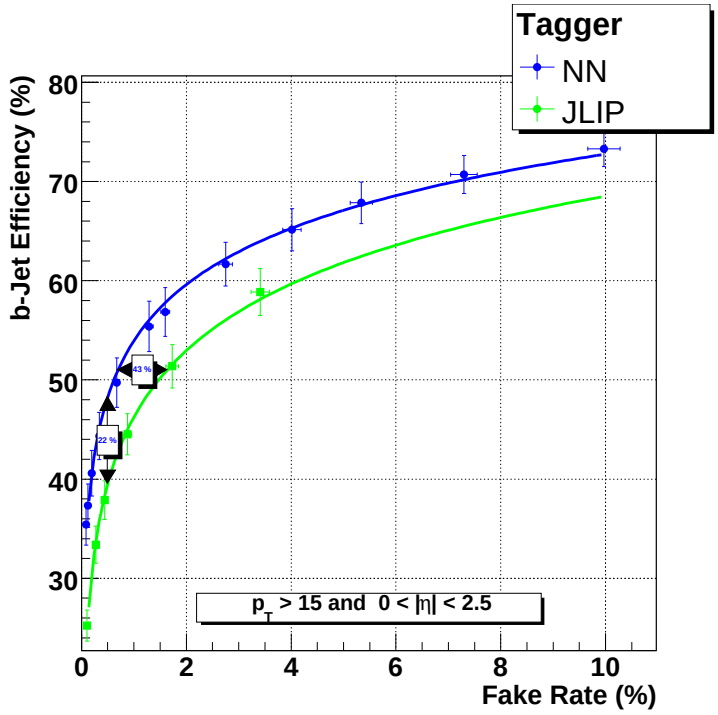


(g) the number of SV found within the jet.

Figure 4.16: Input variables to the NN tagger [139].



(a)



(b)

Figure 4.17: Figure (a) shows the output of the NN tagger, while figure (b) shows the performance of the NN tagger for jets with $p_T > 15\,\text{GeV}$ and $|\eta| < 2.5$ [139].

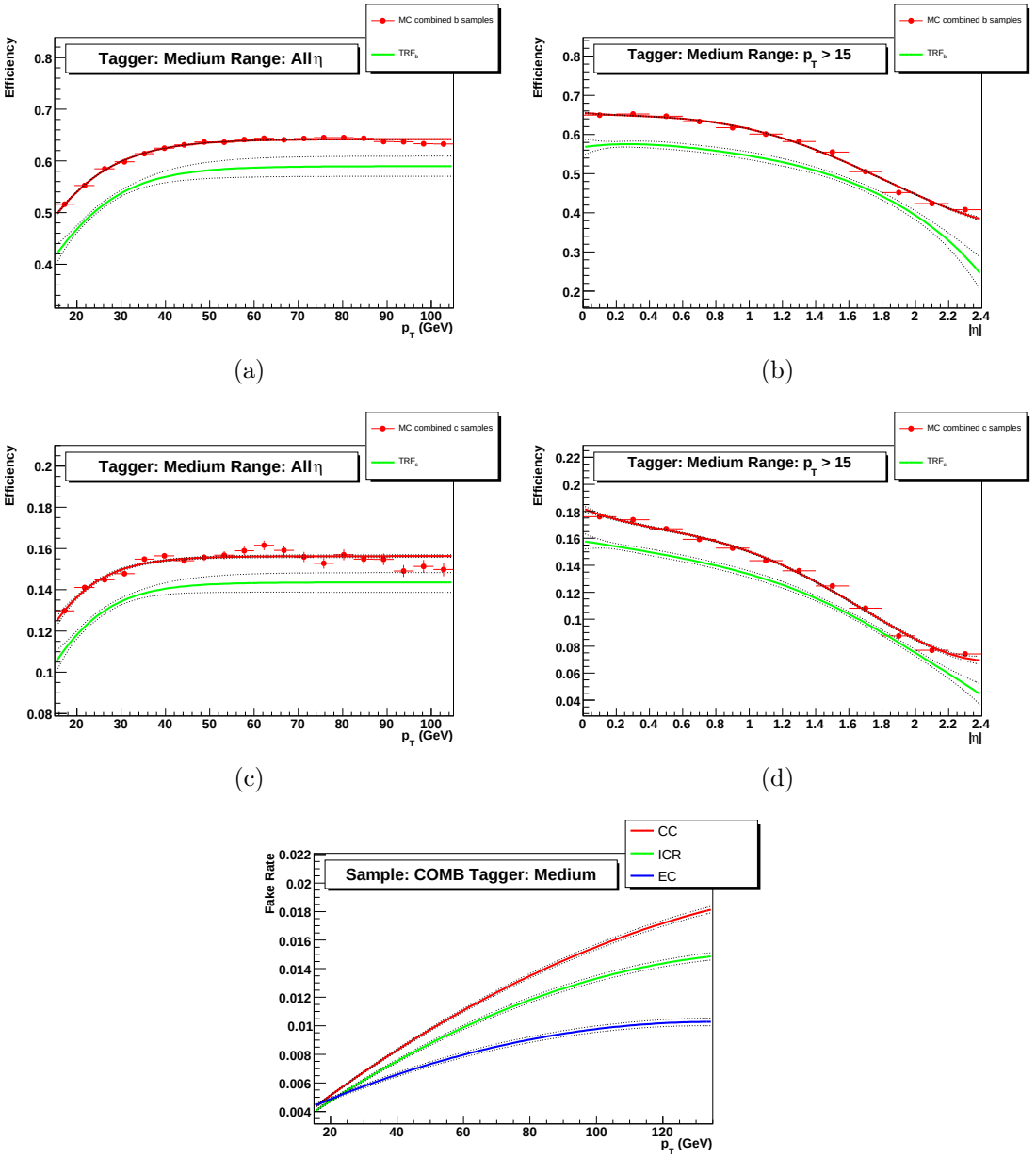


Figure 4.18: *b*-, *c*- and light-TRF for the medium NN tagger [139].
Figure (a) shows the *b*-tag efficiency on simulated events and *b*-TRF parametrized in p_{T} and (b) for η. The *c*-tag efficiency on Monte Carlo and *c*-TRF parametrized in p_{T} is shown in figure (c) and for η (right) in figure (d). Figure (e) shows the fake tag rate parametrized in p_{T} in the three η regions CC, ICD, and EC.

4.5 Missing Transverse Energy

Neutrinos only interact through the weak interaction and therefore leave the detector without a trace. In the transverse plane the momentum conservation requires the total E_{T} to be zero. The missing transverse energy $\not\!\!E_{\mathrm{T}}$ can therefore be used to identify neutrinos. The x and y components of the $\not\!\!E_{\mathrm{T}}$ are given by $\not\!\!E_{\mathrm{x}} = -p_{\mathrm{x}}^{\mathrm{meas}}$ and $\not\!\!E_{\mathrm{y}} = -p_{\mathrm{y}}^{\mathrm{meas}}$ [119, 143]. $p_{\mathrm{x}}^{\mathrm{meas}}$ and $p_{\mathrm{y}}^{\mathrm{meas}}$ are the vectorial summed momenta measured in the electromagnetic and fine hadronic calorimeter, ICD and massless gap cells after the removal of noise with the T42 algorithm [144]. If the event has an identified muon, the $\not\!\!E_{\mathrm{T}}$ has to be corrected for the energy deposited by the muon in the calorimeter, which must be subtracted. The $\not\!\!E_{\mathrm{T}}$ is corrected for coarse hadronic calorimeter cells, which are not used in the calculation due to their high noise ratio. The $\not\!\!E_{\mathrm{T}}$ needs to be corrected for the JES correction as it changes the balance of the jets in the transverse plane. The correction therefore is applied before the summation of the objects to calculate $\not\!\!E_{\mathrm{T}}$. For this analysis, the $\not\!\!E_{\mathrm{T}}$ is further changed to be only applied on the objects used for the $t\bar{t}$ event reconstruction. Finally, the $\not\!\!E_{\mathrm{T}}$ is then calculated by $\not\!\!E_{\mathrm{T}} = |\not\!\!E_{\mathrm{T}}|$.

CHAPTER 5

Properties of Data Samples and Monte Carlo Simulation

This chapter provides a summary of the observed data and its quality requirements as well as the used Monte Carlo simulations as well as their corresponding generators. The observed data, which consist of about $3.6\,\mathrm{fb}^{-1}$ in both of the decay channels needs to fulfill quality criteria to ensure good data quality as described in the first section. A detailed explanation of the data quality requirements and used trigger lists can be found in reference [141, 145]. The Monte Carlo simulation, described in the second section, is passed to a detector simulation and finally reconstructed in the same way as the observed events. The last section of this chapter will describe the different Monte Carlo sample corrections and weights used within DØ, to give an insight to judge the effect on the sample and their meaning for the systematic uncertainties described in chapter 6.

5.1 Observed Data

The analyzed DØ data set includes about $1.0\,\mathrm{fb}^{-1}$ from Tevatron Run IIa, which was taken from August 2002 to December 2005, and about $2.6\,\mathrm{fb}^{-1}$ from Run IIb, taken after the detector upgrade from January 2006 to July 2008. While Run IIb is still ongoing, the Run IIa data set is complete. For all Run IIa data and Monte Carlo samples the software release p17 was used, while all Run IIb data and Monte Carlo samples used software release p21 for reconstruction.

5.1.1 Data Quality

Measured events need to be qualified as good, to be used in this analysis. Detector components might fail during single events or even through longer periods. The data is therefore divided into runs which are in turn subdivided into luminosity blocks as discussed in chapter 3. If one of the detector subcomponents or even more is marked as bad, the luminosity block or even the whole run is marked bad. The conditions for a specific run are stored in the Offline Run Quality Database. The entries are filled by the data taking shifters, marking components of the detector that are not working properly during a run.

Entire luminosity blocks are declared bad based on calorimeter information collected in each of the blocks. Multiple possibilities exist to qualify a block as bad. Examples are the so called "ring of fire", where grounding problems give a ring of energy in ϕ, "coherent noise", where all readout crates experience a synchronous shift in the pedestals, or long lasting missing transverse energy measurements significantly different from zero. If more than 20 consecutive luminosity blocks show this behavior, all of those blocks are discarded.

In case of event based flags, those events are removed from the analysis but cannot be easily removed from the luminosity calculation. Non event based flags are covered by making an efficiency correction for the Monte Carlo simulation.

5.1.2 Trigger

Hadron colliders produce events at a rate, which is by far larger, due to the frequency and size of the events, than can be handled by even the most modern data acquisition system. Therefore a subset of the data events needs to be chosen in a fast way. This method is called triggering and the algorithms in use are called triggers. A set of triggers used simultaneously is called trigger list. This analysis uses a variety of trigger lists. Table 5.1 gives the corresponding integrated luminosity acquired during the different run periods.

The trigger used for the Run IIa and RunIIb data sets require either at least one jet and an electron like object for the e+jets channel or for the μ+jets channel a muon like object plus at least one jet. The triggers are combined to a so called SuperOr trigger. If at least one of the triggers has fired in the corresponding lepton channel, the event will be stored. The efficiency of these triggers is in sum near is 100 %, which makes them standard for all ongoing and coming analysis. A more detailed description can be found in [146, 147, 148].

The Run IIa and Run IIb data set together have an integrated luminosity of about $3.6\,\mathrm{fb}^{-1}$.

Run Period	Sample	Integrated Luminosity [pb^{-1}]
Run IIa	e+jets	1037.82
	μ+jets	996.27
Run IIb	l+jets	2580.08

Table 5.1: Integrated luminosity collected with the e+jets trigger and the triggerlist version for the full Run IIa data set.

5.2 Monte Carlo Simulation

For this analysis, a detailed knowledge of the background contributions is needed to differentiate between the SM and possible beyond SM contributions. Monte Carlo generators are used to simulate the various physical processes. The simulation process can usually be divided in several parts.

The process is ordered from high energy to low energy processes or from short to long timescales, where long is still well below microseconds. The first part usually calculated is the hard matrix element, where the high momentum transfer takes place. It describes the physical process, the Monte Carlo simulation is accounted to. This is essentially a transition from a initial state to a final state, where the final state is usually chosen to produce only events with the desired structure, like for example $t\bar{t}$ pair production. The process can be described by Feynman diagrams as seen in chapter 2.

At hadron colliders, this involves the determination of the momentum fraction x each parton has and the energy scale Q^2, which is the basically the transferred momentum in the process. This information is parametrized in PDFs, which tell the probability to find a parton with the desired values in x and Q^2. The PDFs are calculated for leading order matrix elements in this thesis, though next to leading order PDFs exist as well.

Once the matrix element is calculated the outgoing particles decay. Quarks emit gluons or may in case of heavy quarks decay weakly, if the lifetime permits this. Gluons split into two or three gluons or produce quark antiquark pairs, similar to photons. All electromagnetic charged objects emit photons as well. The energy scale, at which the hard matrix element processing stops and the parton shower evolution starts varies depending on the parameters set within the Monte Carlo generator. In principal this should have no effect on the physics observed.

Additionally radiation before and after the parton collision, so called Initial State Radiation (ISR) and Final State Radiation (FSR) needs to be accounted

for. Initial state radiation is done before the hard matrix element, making the radiated parton the new initial parton. Final state radiation is in principal coped within the parton shower or the hard matrix element depending on the energy scale and generator chosen for this.

At this point all particles may have a color charge, while colored objects are not observed in nature. A process called hadronization or fragmentation takes place, which is not well understood in theory. The process models the transfer of colored partons to color neutral hadrons. This process needs to obey the conservation laws but is a rather phenomenological model.

Once this step has been processed, pile-up may be added to simulate the high luminosities found in hadron colliders. This can be done by adding more random Monte Carlo simulated events or by adding real data.

After this a detector simulation is added, propagating all particles through a virtual detector and simulating the passing through matter. This step can be left out, if no detector is specified but this is rare.

If the detector was simulated, then the response of the detector electronics is simulated. This involves for example electronic noise and resolution effects. The simulated events are passed through a reconstruction process which is in an optimal case identical to the one used for real events. Finally, the simulated events are treated like collision events plus having the MC true information.

5.2.1 DØ Handling of Simulated Events

All simulated events are processed through the full Run IIa / Run IIb detector simulation and event reconstruction. The DØ detector is simulated by `D0gstar`, which uses `GEANT` [149, 150, 151] for the simulation of the DØ detector. This program simulates the behavior of particles traveling through matter. Its output is passed to `D0sim` [152], that simulates the digitization of the electronic signals produced by `D0gstar` including the pile-up from previous bunch crossings as well as the detector noise and multiple proton proton interactions. The last two could be accounted for by Monte Carlo simulations, but the program uses real minimum bias data, which is recorded without trigger requirement. These events are sorted according to their instantaneous luminosity. During the MC generation events from a matching instantaneous luminosity required for the desired MC simulation are drawn at random from the pool of recorded minimum bias data and added to the MC simulation. This simulates a random event including all detector noise from the real detector, which is better than the simulation.

A program called `D0reco` which is used for data and Monte Carlo simulations alike reconstructs the events. The only difference between observed data events

and simulated events is that the latter also carry the Monte Carlo true information.

Triggers are not simulated in the Monte Carlo simulation and are therefore measured in data and parametrized as a function of lepton and jet p_{T}, η or ϕ. This parametrization is applied to Monte Carlo events and a probability to fire a trigger is assigned to each event.

The DØ Monte Carlo web page offers a more detailed information [153].

5.2.2 Monte Carlo Samples

All Monte Carlo samples have been generated with the CTEQ6L1 PDF set [154]. For the Run IIa samples PYTHIA 6.323 was used, while for the Run IIb samples PYTHIA 6.409 was used, unless noted otherwise. Jet Shifting and Jet Removal (JSSR), see section 4.3.4 were turned on for W+jets, Z+jets, WW, WZ, ZZ and single top backgrounds. The $t\bar{t}$ background is supposed to have the correct modeling of quark vs. gluon jets in the Monte Carlo simulation and therefore the JSSR is not turned on. This was studied by the top group and made the common analysis standard. A discussion of these cuts can be found in Ref. [155, 156].

The resonance signal Monte Carlo is generated with PYTHIA for ten different resonance masses M_X: 350, 400, 450, 500, 550, 600, 650, 750, 850, and 1000 GeV using $m_t = 175\,\mathrm{GeV}$ and $m_t = 170\,\mathrm{GeV}$ for Run IIa and Run IIb, respectively. The width of the resonance is set to $\Gamma_X = 0.012 M_X$. This qualifies X as a narrow resonance, since its width is smaller than the mass resolution of the DØ detector of about 10 %. The exact value of the width of 1.2% was chosen in order to compare the results with former Run IIa and CDF analyses where this width was used. The resonance particle is forced to decay to t and $\bar{t}$ only and the decays are inclusive. A list of the main parameters used in PYTHIA is provided in appendix B. Shape comparisons between $t\bar{t}$ and the different resonance samples can be found in section 6.4.2.

The SM $t\bar{t}$ process, the W+jets, and Z+jets samples are generated with ALPGEN [157] for the hard interaction plus PYTHIA for the parton shower, hadronization, and hadron decays using a top quark mass of 175 GeV (172.5 GeV) for Run IIa (Run IIb). The factorization scale was set to $m_t^2 + \sum p_T^2(jets)$. A top mass of 175 GeV and 172.5 GeV was used for Run IIa and Run IIb. To avoid double counting between the hard matrix element and the parton shower, the so called MLM jet-matching algorithm was applied [158] here. Double counting occurs, when event generation is performed with matrix elements including various numbers of jets followed by a parton shower. Because parton showers may add additional jets to the hard event, certain configurations may be produced either

by a matrix element with an additional jet from the parton shower, or directly from that matrix element, that includes the additional jet in the first place. The semileptonic $t\bar{t}$ decays include the τ+jets channel. For the τ decays a parton filter was applied to select only the leptonically decaying τs.

The diboson background WW, WZ and ZZ samples were generated with PYTHIA. Single top production for both s- and t-channel was generated using the COMPHEP single top generator [159].

The MC samples for $t\bar{t}$, WW, WZ, ZZ, single top, and Z+jets are normalized to theoretical NLO or NNLO cross sections.

5.3 Corrections and Weights to Simulated Events

Some distributions obtained from simulated events still show some discrepancies to the measured distributions. There are multiple reasons for this, for example leading order calculations for the matrix elements that give a wrong cross sections and sometimes shapes of the simulated events, depending on the generator used, imperfect tracking simulation in the detector, and trigger efficiencies.

These effects change the distributions used and need to be corrected. Though some of the imperfections require a complete reprocessing of the complete Monte Carlo generation scheme, some of them can be corrected by reweighting events or single variables in the event. A brief overview over the reweightings is presented in the following to give an insight into the production as well as the understanding of the systematic uncertainties used later in this analysis.

- The total amount of W+jets events is determined by subtracting all simulated background events, including multijet QCD from the data events before b-tagging. The remaining number of events is used for the normalization of the W+jets background. The shape is taken from the Monte Carlo simulation. The simulated events do not have the correct amount of heavy flavor events, which needs to be adjusted by the W+jets heavy flavor scale factor to the amount of the W+jets heavy flavor fraction measured in data. The scale factor can be determined by [128]:

$$k_B = \frac{(data' - X') \cdot W'' - (data'' - X'') \cdot W''}{data'' - X'' \cdot B' - (data' - X') \cdot B''}, \tag{5.1}$$

  where k_B is the k factor for the Wbb+jets sample, X the sum of all simulated events except the W+jets sample and W (B) the number of Wlp+jets (Wbb+jets) events, where lp denotes light partons. The b-tagged sample is marked by $'$ while the anti-b-tagged sample is marked by $''$. A more detailed description is available at [128].

- The tracking system in the detector description is not modeled well enough to describe the real data. To account for the in general higher probability to reconstruct, identify, and select leptons in simulated events a correction factor is applied as a function of η_{det} and ϕ on simulated events. The correction factors are measured in observed data and range from $\sim 92\,\%$ to $\sim 101\,\%$ [102, 160].
- The Monte Carlo samples generated with ALPGEN or PYTHIA do use leading order (LO) matrix elements. Events generated with ALPGEN are additionally split up for different parton multiplicities. This is also useful to enrich statistics for higher jet multiplicities. These get a weight assigned corresponding to the leading order cross section expected for the separate parton multiplicity. To account for the next to leading order (NLO) calculations an overall k-factor is applied on the event weight to account for the difference between LO and NLO calculations. The factor is determined individually for each sample.
- The data quality depends on the functionality of the calorimeter and as mentioned in the luminosity calculation sometimes single events are taken out of the observed data due to malfunctions. These cannot be calculated by removing luminosity blocks. Therefore a correction factor for the selection efficiencies on the simulated events have to be taken into account which was determined to be $97.14 \pm 0.3\,\%$ [161]. This was done by triggering zero bias data, which is expected to be unbiased by calorimeter noise. A systematic uncertainty of $0.5\,\%$ is assigned to take into account the differences between the considered data samples.
- Simulated events generated for DØ are overlaid with minimum bias interactions which are recorded with the instantaneous luminosity of the collider at that time. As the instantaneous luminosity increased over the time of data taking, the luminosity profile of the minimum bias events does no longer match the profile of the current data events. As discussed before the minimum bias events simulate detector and luminosity effects, especially pile-up. The Monte Carlo samples are reweighted in a way that the luminosity profile of the observed data matches the one in the simulated events. The luminosity profile is chosen corresponding to the data sample used. Depending on the applied Monte Carlo sample the effect can be up to a few percent.
- The z position of the PV is modeled as a Gaussian with 25 cm width and the center of the detector in simulated events. In real data, the position and shape varies depending on several factors. To correct the z distribution in the simulated events a weight is applied to match the shape in data [162]. This is correlated to the luminosity reweighting before, as the shape also depends on the instantaneous luminosity. Nevertheless, the former alone

does not model the observed data in a precise way.

- As mentioned before the triggers are not modeled in simulated events in the same way as in observed data. The efficiencies for the triggers are measured in data and parametrized to a probability for firing a certain trigger [163].
- The reweighting of the b-fragmentation is needed, as the fraction of the energy plus longitudinal momentum (z) taken by the heavy hadron is not the same as measured in data. The B hadrons coming from the hadronization process therefore get reweighted to match the tuning on data [164]. This has a small effect on the b-tagging efficiency.
- The Z boson p_{T} differs in observed Z boson enriched data samples and simulated Z+jets samples. Therefore a reweighting procedure was developed [165] to correct the p_{T} of the Z boson. As the contribution is small in the l+jets channel, it does not affect the results as opposed to analyses performed in the dileptonic decay channel.
- Further reweightings include JES, JER, JSSR, TaggabilityRFs, TRFs, ϵ_{sig}, and ϵ_{QCD}, which were described in the previous chapter.

CHAPTER 6

Data Analysis

This chapter focuses on the analysis. First, the selection of the events which enhance the signal-to-background ratio is discussed. The physical reason for the applied cuts is the main topic of discussion. Secondly, the treatment for the different contributions of background is explained. Afterwards the reconstruction of the $t\bar{t}$-invariant mass distribution is discussed, where the change to a direct reconstruction method in comparison to the kinematic fit used in previous analyses is explained. Then the systematic uncertainties taken into account are explained. A discussion of the statistical method used in the analysis concludes the chapter.

6.1 Selection

The used cuts aim to filter semi-leptonic top pair events and to optimize the signal-to-background ratio, while not loosing too many events in the process. The final state in a semileptonic $t\bar{t}$ decay consists of four quarks, with at least two b quarks, one lepton, and one neutrino. Ideally, the four quarks are detected as four jets in the detector, as well as an isolated electron or muon for the lepton and the missing transverse energy for the undetected neutrino. Due to the high mass of the top quark and the boost of the $t\bar{t}$ pair, the W boson and the b quark may both have large momenta. The W boson decays into two light quarks with almost no rest mass compared to the boson itself. A so called W boson monojet, where both quarks are reconstructed as one jet, may appear, if the boost is large enough. For higher boosts even the t quark may decay into a top quark monojet by combining three final state quarks, two of which are from the W boson and one from the b quark, into one jet. This changes the signature of the event in

the detector to three or two jets. For the center of mass energy available at the Tevatron collider, the two jet case is a rare case, though a significant number of $t\bar{t}$ events end up with three reconstructed jets, as can be seen in figure 7.1. The total loss of a jet due to acceptance or jet energy loss is more likely than top monojets. To resemble the final state of the $t\bar{t}$ decay, a high p_T tight electron or an isolated muon in the e+jets and μ+jets channel, respectively, at least three jets and high missing E_T is required. The same cuts were applied measuring the $t\bar{t}$ cross section in the DØ-experiment, where at least four jets were requested in contrast to this analysis.

Before applying the final selection, a preselection is made to select the important subset of all data events. Later on, the analysis presented uses the b-tagging to split the preselected dataset into subsamples, which are either used to extract the result or serve as a control samples for the correct modeling of the Monte Carlo simulations.

6.1.1 Common Selection

The preselection for this analysis is based on the $t\bar{t}$ cross section measurement [145] and has been optimized for $t\bar{t}$ selection. The resonances studied in this analysis decay into a $t\bar{t}$ pair and therefore have the same final state as the SM $t\bar{t}$ selection, which allows to apply the same criteria.

As described in the last chapter the datasets are split up in the Run IIa and Run IIb dataset. Wherever the parameters for the selection are different for the runs the values are stated separately. The correction factors used are derived independently and are applied only to their appropriate part, i.e. Run IIa JES is only applied to Run IIa data and Monte Carlo simulation, and Run IIb JES only to Run IIb data and Monte Carlo simulation.

The e+jets and μ+jets selections share a number of cuts. The cuts for the separate channels will be handled in the next section. Since the analysis is using the semileptonic channel, a second lepton is vetoed in the analysis. The common selection shares the following requirements:

- Good data quality as described in the section 5.1.1.
  The good data quality flag ensures, that all used detector systems are fully functional. For this analysis this includes practically every subsystem from the inner detector for identifying b-quark jets, to the calorimeter for jet and electron reconstruction to the muon system.
- At least three jets with $p_T > 20\,\text{GeV}$, $|\eta| < 2.5$.
  This ensures the inclusion of the three and four or more jet signatures. The p_T cut is imposed as jets with the minimal possible transverse momentum

of 7 GeV up to about 15-20 GeV have a large uncertainty on the JES. The region which is cut out, is also dominated by multijet events, which further enhances the signal-over-background ratio. Apart from this only the central η region is used, where most of the $t\bar{t}$ events are expected and the calorimeter has a good performance.

- For Run IIb vertex confirmed jets are required.
  This means, that at least two tracks from the primary vertex can be attached to the jet. This is not needed for Run IIa data, mainly because the instantaneous luminosity was lower and jets from pile-up events were not as common. Apart from that the additional layer in the silicon detector enables a quality cut without loosing much signal.
- A leading jet with $p_T > 40\,\text{GeV}$.
  This cut suppresses a major amount of the W+jets background, loosing only a minor amount of signal, as only few W+jet events have a p_T higher than 40 GeV, while a jet from a top quark decay has a higher p_T.
- A good vertex within $|z_\text{PV}| \leq 60\,\text{cm}$ with at least three tracks attached.
  This ensures that b jets can be properly identified and the vertex position is well localized.

Only events fulfilling the above cuts are used in the principal measurement of the result. Events with one or two jets matching the rest of the preceding criteria are kept for control plots. Jet bins are calculated from jets that fulfill the above cuts and after the JES corrections as well as the corrections for observed muons are applied.

6.1.2 Electron + Jets Channel

In addition to the common selection, the electron channel requires:

- One tight electron [166] with $p_\text{T} > 20\,\text{GeV}$ in the CC ($|\eta| < 1.1$).
  As electrons produce a shorter shower than jets, only the CC is used to minimize the chance of a missed electron in the ICD region. The p_T cut ensures that the trigger turn on for the combined trigger is flat and the electron is properly selected.
- The electron must come from the PV ($|\Delta z(e, PV)| < 1$ cm).
  This suppresses electrons from secondary decays, underlying events, or pile-up events.
- Veto on a second tight electron with $p_\text{T} > 15\,\text{GeV}$ in the CC or EC.
  This cut ensures the orthogonality to the dilepton selection, as well as the reduction of $Z \rightarrow ee$ events. The cut energy is reduced to have a better rejection on the events.

- Veto on a tight muon with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, $|\eta| < 2.0$.
  Most important, this ensures the orthogonality to the μ+jets selection and also to the dilepton selection. The η region differs compared to the electrons, as the muon system has a larger central region.
- Missing transverse energy $\not\!\!E_{\mathrm{T}} > 20\,\mathrm{GeV}$
  This indicates a neutrino in the event, while a lower threshold might cause an event to be taken, where an underlying event with objects at high η give a nonzero transverse momentum for the sum of all objects.
- Triangle cut of $0.7\pi - 0.045\,\not\!\!E_{\mathrm{T}} < \Delta\phi(\vec{e}, \vec{\not\!\!E}_{\mathrm{T}})$
  This cuts a region in parameter space, where we have low $\not\!\!E_{\mathrm{T}}$ and a small polar angle between the isolated lepton and the $\not\!\!E_{\mathrm{T}}$. The triangle cut rejects multijet events where a jet is misreconstructed as an isolated lepton, which leads to the conditions discussed before.

6.1.3 Muon + Jets Channel

In addition to the common selection the following criteria are required in the muon channel:

- One tight muon with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, $|\eta| < 2.0$.
  For muons the full central region is used which limits the η range to 2.0. A sufficiently large p_{T} for the muon is required to have muons from the PV.
- Invariant mass of muon and any second muon $m_{\mu\mu} < 70\,\mathrm{GeV}$ or $m_{\mu\mu} > 100\,\mathrm{GeV}$.
  This excludes events which are close to the Z boson mass peak and therefore suppresses $Z \to \mu\mu$ events. The second muon needs to be central ($|\eta| < 2.0$) and has to have a p_{T} of at least 15 GeV.
- No loose medium muon with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ with muon quality MediumNSeg3.
  This further enhances the $Z \to \mu\mu$ rejection as well as ensuring orthogonality to the dilepton selection.
- No tight electron with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ in the CC.
  As before, this excludes dilepton events and orthogonality to the e+jets channel.
- Prompt muon, coming from the primary vertex: $|\Delta z(\mu, \mathrm{PV})| < 1\,\mathrm{cm}$.
  This excludes muons from secondary decays, underlying events, or pile-up events.
- Missing transverse energy $\not\!\!E_{\mathrm{T}} > 25\,\mathrm{GeV}$.
  As above, this indicates a neutrino in the event.

- Triangle cut of $2.1 - 0.035\not{E}_{\mathrm{T}} < \Delta\phi(\vec{\mu}, \vec{\not{E}}_{\mathrm{T}})$.
  Again this cut rejects the multijet events.

Compared to the Run IIa data, vertex confirmed jets were used to give a better description of the jet p_T's in Run IIb. This was studied by the top group and was part of the common analysis standard. A discussion of these cuts can be found in Ref. [155, 145, 156]. Control plots for the preselection before and after b-tagging using the medium working point of the NN tagger for Standard Model prediction can be found in Appendix C.

6.2 Background Estimation

The two main backgrounds, besides $t\bar{t}$ pair production itself, to the resonance production in the l+jets channel, are the electroweak W boson production accompanied by multiple jets (W+jets) and QCD multijet events, where a jet fakes an electron or a muon. Aditionally, $t\bar{t} \rightarrow l\bar{l}$, single top production, diboson (WW,WZ and ZZ), as well as Z/γ^* production are considered as background sources in this analysis.

The estimation of the number of expected events for the different backgrounds differs depending on the source. First, the QCD multijet background is estimated from observed events. They model best this instrumental background, where jets with a high electromagnetic fraction fake electrons or heavy flavor decays fake PV muons. Further backgrounds, where the final state is similar to resonant $t\bar{t}$ pair production are described. This includes $t\bar{t}$ pair production in the semileptonic and dileptonic channel, single top production, diboson (WW, WZ, ZZ) production and Z+jets production. W+jets production though a physical background is treated slightly differently, as described in the following.

All backgrounds and signal samples use the PDF set CTEQ6L1. MC-to-data correction factors [141] are applied to the background samples as well as to the resonance signal samples. The number of expected events for all background samples but the W+jets and multijet background is calculated accordingly to their cross section, selection efficiency, and the luminosity from the observed number of events.

6.2.1 Estimation of Multijet Background

The overall normalization of the multijet background is determined from lepton triggered events using the following method, also called matrix method. Two samples of events, a loose isolated and a tight isolated set, where the latter is a

subset of the first, are needed. The tight sample (N_t), which is the candidate sample for the analysis, requires a tight muon isolation in the μ+jets channel and a likelihood cut in the e+jets channel. The loose set (N_ℓ) corresponds to the preselected sample without the tight lepton cut, i.e. requiring a loose lepton selection as stated in [145]. Using the number of events with a real lepton (W+jets, top, diboson or Z+jets), $N_{W+\text{top}}$, and the number of events with a fake lepton, N_{QCD}, the number of loose and tight events can be written as:

$$\begin{aligned} N_\ell &= N_{W+\text{top}} + N_{\text{QCD}} \\ N_t &= \varepsilon_{\text{sig}} N_{W+\text{top}} + \varepsilon_{\text{QCD}} N_{\text{QCD}} \ . \end{aligned} \tag{6.1}$$

Here ε_{sig} is the efficiency for a true isolated lepton to pass the tight isolation cuts, while ε_{QCD} is the fraction of multijet events passing the tight isolation criteria.

The efficiency ε_{QCD} is taken from data [167] by taking events with the standard selection except requiring $\not\!\!E_{\text{T}} < 10\,\text{GeV}$. This sample is enriched in multijet QCD events. No real leptons are assumed to be included in this sample, as the contamination with leptons is a few percent. The ratio of events with a tight isolated lepton over a loose isolated lepton gives the fake rate or ε_{QCD}.

The efficiency ε_{sig} is taken from Monte Carlo simulations and the mean of the W+jets and the $t\bar{t}$ samples. As the event kinematics differ in both, the difference is assigned as a systematic uncertainty to ε_{sig}.

Solving this linear system for N_{QCD} and $N_{W+\text{top}}$ yields:

$$N_{W+\text{top}} = \frac{N_t - \varepsilon_{\text{QCD}} N_\ell}{\varepsilon_{\text{sig}} - \varepsilon_{\text{QCD}}} \quad \text{and} \quad N_{\text{QCD}} = \frac{\varepsilon_{\text{sig}} N_\ell - N_t}{\varepsilon_{\text{sig}} - \varepsilon_{\text{QCD}}} \tag{6.2}$$

The values for ε_{QCD} and ε_{sig} are shown in table 6.1. The procedure to derive them is described in detail in [141]. The number of loose and tight events can be found in table 6.2. The predicted number of multijet events obtained from this method for the various channels is listed in line 2 of tables 6.3–6.8 for Run IIa and tables 6.9–6.14 for Run IIb.

The shape of the QCD background obtained from a bin by bin matrix method suffers from low statistics. Therefore only the total number of QCD events within each jet bin is taken from the matrix method and the shape is derived from QCD events passing the loose, but not the tight preselection.

The non-QCD contamination in the loose sample is lower than 3 % in Run IIa and Run IIb. The determination of the uncertainty on the multijet background is discussed in detail in [168].

		e+jets	μ+jets
p17	ϵ_{QCD}	0.19 ± 0.02(stat+syst)	0.28 ± 0.05(stat+syst)
	ϵ_{sig} ($= 1$ jet)	0.84 ± 0.04(stat+syst)	0.91 ± 0.01(stat+syst)
	ϵ_{sig} ($= 2$ jets)	0.85 ± 0.02(stat+syst)	0.89 ± 0.01(stat+syst)
	ϵ_{sig} ($= 3$ jets)	0.85 ± 0.01(stat+syst)	0.87 ± 0.01(stat+syst)
	ϵ_{sig} (≥ 4 jets)	0.84 ± 0.02(stat+syst)	0.85 ± 0.02(stat+syst)
p20	ϵ_{QCD}	0.13 ± 0.02(stat+syst)	0.09 ± 0.02(stat+syst)
	ϵ_{sig} ($= 1$ jet)	0.82 ± 0.01(stat+syst)	0.87 ± 0.01(stat+syst)
	ϵ_{sig} ($= 2$ jets)	0.81 ± 0.01(stat+syst)	0.84 ± 0.01(stat+syst)
	ϵ_{sig} ($= 3$ jets)	0.80 ± 0.01(stat+syst)	0.81 ± 0.01(stat+syst)
	ϵ_{sig} (≥ 4 jets)	0.80 ± 0.02(stat+syst)	0.80 ± 0.02(stat+syst)

Table 6.1: Efficiencies for the tight selection in both channels. Values have been rounded to two decimals.

6.2.2 Top Quark Background

The expected number of events is calculated by assuming a Standard Model $t\bar{t}$ cross section of $7.48^{+0.56}_{-0.72}$ pb for the $t\bar{t} \rightarrow l$+jets and $t\bar{t} \rightarrow l\bar{l}$ backgrounds. This represents the currently best NNLO approximation on the $t\bar{t}$ cross section for a top quark mass of 172.4 GeV [37]. We use this cross section to compute the expected $t\bar{t}$ contribution independent of the top quark mass used for kinematics in the generation of events.

6.2.3 W+jets Background

The contribution of W+jets events before b-tagging is calculated by substracting all other Monte Carlo backgrounds and the multijet contribution, from the observed number of events. W+Jet events consist of Wbb+jet, Wcc+jet, and Wlq+jet events, with lq denoting light quarks. The relative contribution of these three subclasses cannot purely be taken from the cross section generated by ALPGEN, as it does not take into account next-to-leading order corrections. Additionally to the relative fraction determined from Monte Carlo. The heavy flavor contributions Wbb+Jet and Wcc+Jet are multiplied by a heavy flavor scale factor of 1.93 for Run IIa and 1.43 for Run IIb. This difference is due to the different ALPGEN versions used during both run periods. The uncertainty on the flavor scale factor is about 20%. The number of W+jets events after b-tagging is calculated as the product of the number of untagged W+jets events and the probability $P^{\text{tag}}_{\text{event}}$, to have a tag in the event computed from the TRFs.

	Channel		before b-tagging		with 1 b-tag		with 2 b-tags	
			N_ℓ	N_t	N_ℓ	N_t	N_ℓ	N_t
Run IIa	3-jets	e+jets	2330	1602	350	192	63	45
		μ+jets	1712	1360	191	147	45	30
	$\geq$ 4-jets	e+jets	827	419	168	115	40	30
		μ+jets	496	385	124	106	37	35
Run IIb	3-jets	e+jets	7693	3692	818	470	154	109
		μ+jets	3639	2829	433	338	85	72
	$\geq$ 4-jets	e+jets	2053	986	382	244	118	88
		μ+jets	1093	826	271	217	113	99

Table 6.2: Observed data events for the loose and tight preselection before and after b-tagging as input for the matrix method. Numbers shown for the b-tagged samples correspond to the medium working point of the NN tagger.

6.2.4 Z+jets and Electroweak Backgrounds

To estimate the physics background due to different electroweak processes, the number of events after preselection is estimated from Monte Carlo calculations, taking into account the Monte Carlo selection efficiencies and b-tagging efficiencies.

- Diboson background: WW, WZ, and ZZ samples generated with PYTHIA are used. The PDF set is CTEQ6L1. The NLO cross section [119] of 12.0, 3.68 and 1.42 pb are taken respectively. According to [169], the uncertainty on the cross section is estimated to be 6.8% for all diboson samples.

- Single top-quark background: The single top-quark contribution from the s- and t-channel are used. The samples are generated with the COMPHEP single top quark Monte Carlo generator [159]. The cross section is 1.04 and 2.26 pb with 12.5% uncertainty [170, 171]. The top quark mass is set to 172.5 GeV.

- Z+Jets background: All different available Z+Jets [172] samples simulated with ALPGEN [157] are taken into account. The Z boson p_{T} reweighting described in [172] is applied, to increase the data/MC agreement of the Z boson p_{T} distribution. To determine the contribution, the Monte Carlo cross section is multiplied with a scale factor of 1.30 for light Z+Jets Monte Carlo and 1.67 for $Z + b\bar{b}$, and 1.52 for $Z + c\bar{c}$ [173]. On the cross section uncertainties of 15% for both the light Z+Jets and heavy flavor Z+Jets are used [174].

6.3 Identification of b-Jets / Flavor Tagging

The usage of neural net information to identify b-jets improves the signal to background ratio. The description of the NN tagger and its performance is given in section 4.4.2 as well as in [139, 140].

In this analysis two separate NN working points can be used to select single and double tagged events. Double tagged events are chosen with a less stringent NN working point (L). From those events which are not selected as doubly tagged, a possibly more stringent NN working point (T) is applied to separate singly tagged events from no-tag events.

The no-tag, single and double tag efficiencies are determined by evaluating the per jet tag efficiency $P_f^W(E_T, \eta)$ for a given flavor (f) and working point ($W = L, T$) [175] and multiplying the efficiencies for all jets in the event:

$$P_{\text{event}}^{\text{tag}}(n=1) = \sum_{j=1}^{N_{\text{jets}}} P_{f_j}^T(E_{Tj}, \eta_j) \prod_{i=1, i\neq j}^{N_{\text{jets}}} (1 - P_{f_i}^L(E_{Ti}, \eta_i)). \tag{6.3}$$

$$P_{\text{event}}^{\text{tag}}(\geq 2L) = \sum_{i,j}^{N_{\text{jets}}} P_{f_i}^L P_{f_j}^L \prod_{k\neq i,j}^{N_{\text{jets}}} (1 - P_j^L) + \sum_{i,j,l}^{N_{\text{jets}}} P_{f_i}^L P_{f_j}^L P_{f_l}^L \prod_{k\neq i,j,l}^{N_{\text{jets}}} (1 - P_{f_j}^L) + \dots \tag{6.4}$$

$$P_{\text{event}}^{\text{tag}}(n=0) = 1 - P_{\text{event}}^{\text{tag}}(n=1) - P_{\text{event}}^{\text{tag}}(n >= 2) \tag{6.5}$$

For Eq. (6.4) terms for up to 5 b-tags are implemented in the computation. Only good jets with $p_T > 20\,\text{GeV}$and $|\eta| < 2.5$ are considered.

Studies made for the Run IIa dataset based on optimizing $S/\sqrt{(S+B)}$ [156] showed that the medium NN working point for both samples is optimal. Therefore, this analysis uses the medium NN working point for selecting both single and ≥ 2 b-tagged jets. Further details are noted in [156].

All expected numbers of events before and after tagging (using the NN medium working point and $n = 1$ or $n \geq 2$ tags) can be found in tables 6.3–6.14. The expected number of events for the different resonance samples are calculated for a cross section of 1 pb. One should note that as the W+jets background is normalized to the observed number of events before b-tagging, the number of total Monte Carlo simulation events is very close to the observed data events.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	$\geq$ 4 Jets
Observed Events	17521	7193	1602	419
$\sum$ Backgrounds	17520.47 ± 218.21	7189.71 ± 44.27	1593.31 ± 13.12	410.74 ± 6.19
Semileptonic $t\bar{t}$	5.80 ± 0.30	54.36 ± 0.85	141.17 ± 1.27	134.21 ± 1.08
Dilepton $t\bar{t}$	7.19 ± 0.11	26.69 ± 0.21	14.06 ± 0.14	4.39 ± 0.06
Single Top	17.64 ± 0.18	43.02 ± 0.28	14.60 ± 0.17	3.96 ± 0.10
Multijet	1144.29 ± 212.53	1055.16 ± 37.99	362.43 ± 10.69	82.78 ± 5.63
$W + lq$+Jets	14880.55 ± 48.39	4746.32 ± 20.88	728.60 ± 6.56	104.85 ± 1.70
$W + bb$+Jets	305.05 ± 2.70	238.46 ± 2.30	57.14 ± 0.96	12.49 ± 0.39
$W + cc$+Jets	689.13 ± 5.38	503.43 ± 4.37	123.63 ± 1.78	27.15 ± 0.73
$Z + lq$+Jets	303.89 ± 8.01	268.32 ± 7.15	78.73 ± 2.73	23.88 ± 1.22
$Z + bb$+Jets	12.82 ± 0.47	15.93 ± 0.51	7.83 ± 0.36	2.52 ± 0.21
$Z + cc$+Jets	23.42 ± 1.43	29.85 ± 1.54	14.68 ± 1.08	4.47 ± 0.43
WW	113.08 ± 0.76	177.08 ± 0.97	41.87 ± 0.48	8.23 ± 0.22
WZ	16.22 ± 0.35	29.36 ± 0.47	7.55 ± 0.24	1.51 ± 0.11
ZZ	1.39 ± 0.06	1.74 ± 0.07	1.03 ± 0.06	0.30 ± 0.03
Z'(m=350 GeV)	0.52 ± 0.07	7.03 ± 0.26	16.18 ± 0.40	17.98 ± 0.43
Z'(m=400 GeV)	0.52 ± 0.07	5.77 ± 0.21	17.12 ± 0.38	20.30 ± 0.41
Z'(m=450 GeV)	0.41 ± 0.05	5.73 ± 0.21	17.51 ± 0.38	22.92 ± 0.44
Z'(m=500 GeV)	0.39 ± 0.06	5.00 ± 0.21	18.05 ± 0.42	24.39 ± 0.49
Z'(m=550 GeV)	0.48 ± 0.05	5.16 ± 0.19	18.54 ± 0.37	26.52 ± 0.44
Z'(m=600 GeV)	0.55 ± 0.06	5.93 ± 0.22	18.43 ± 0.39	26.34 ± 0.47
Z'(m=650 GeV)	0.66 ± 0.07	6.45 ± 0.23	18.70 ± 0.38	25.05 ± 0.45
Z'(m=750 GeV)	1.00 ± 0.09	7.93 ± 0.28	18.73 ± 0.43	22.07 ± 0.47
Z'(m=850 GeV)	1.68 ± 0.11	9.44 ± 0.26	17.72 ± 0.37	20.16 ± 0.41
Z'(m=1000 GeV)	2.75 ± 0.14	10.96 ± 0.29	17.36 ± 0.38	16.70 ± 0.39

Table 6.3: Run IIa predicted and observed number of events in the e+jets channel without b-tag requirements. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	$\geq$ 4 Jets
Observed Events	426	429	192	115
$\sum$ Backgrounds	403.97 ± 6.67	392.58 ± 5.37	176.08 ± 3.34	88.96 ± 2.11
Semileptonic $t\bar{t}$	1.87 ± 0.12	24.34 ± 0.41	64.95 ± 0.60	61.36 ± 0.51
Dilepton $t\bar{t}$	3.38 ± 0.06	12.48 ± 0.10	6.57 ± 0.07	2.02 ± 0.03
Single Top	7.34 ± 0.08	19.70 ± 0.14	6.74 ± 0.08	1.84 ± 0.04
Multijet	38.78 ± 6.41	67.63 ± 5.16	31.09 ± 3.23	7.84 ± 2.03
$W + lq$+Jets	201.51 ± 1.37	96.92 ± 0.66	18.94 ± 0.21	3.33 ± 0.06
$W + bb$+Jets	95.96 ± 1.10	90.52 ± 1.01	21.99 ± 0.42	5.01 ± 0.17
$W + cc$+Jets	42.79 ± 0.50	54.01 ± 0.63	15.25 ± 0.28	4.17 ± 0.13
$Z + lq$+Jets	2.50 ± 0.07	3.82 ± 0.10	1.64 ± 0.06	0.69 ± 0.04
$Z + bb$+Jets	2.91 ± 0.17	5.66 ± 0.21	3.01 ± 0.16	1.05 ± 0.09
$Z + cc$+Jets	1.16 ± 0.11	2.93 ± 0.21	1.78 ± 0.16	0.62 ± 0.07
WW	4.01 ± 0.05	10.84 ± 0.09	3.11 ± 0.06	0.80 ± 0.03
WZ	1.66 ± 0.08	3.53 ± 0.11	0.87 ± 0.05	0.19 ± 0.02
ZZ	0.10 ± 0.01	0.21 ± 0.02	0.14 ± 0.01	0.06 ± 0.01
Z'(m=350 GeV)	0.21 ± 0.04	3.20 ± 0.13	7.52 ± 0.19	8.16 ± 0.20
Z'(m=400 GeV)	0.18 ± 0.03	2.60 ± 0.10	7.89 ± 0.18	9.35 ± 0.19
Z'(m=450 GeV)	0.13 ± 0.02	2.46 ± 0.10	8.10 ± 0.18	10.50 ± 0.20
Z'(m=500 GeV)	0.13 ± 0.02	2.21 ± 0.10	8.38 ± 0.20	11.09 ± 0.23
Z'(m=550 GeV)	0.13 ± 0.02	2.22 ± 0.09	8.51 ± 0.17	12.01 ± 0.20
Z'(m=600 GeV)	0.14 ± 0.02	2.63 ± 0.11	8.50 ± 0.19	12.05 ± 0.22
Z'(m=650 GeV)	0.18 ± 0.02	2.84 ± 0.11	8.66 ± 0.18	11.38 ± 0.21
Z'(m=750 GeV)	0.27 ± 0.03	3.62 ± 0.14	8.63 ± 0.21	10.09 ± 0.22
Z'(m=850 GeV)	0.52 ± 0.04	4.40 ± 0.13	8.13 ± 0.18	9.18 ± 0.19
Z'(m=1000 GeV)	0.96 ± 0.06	5.04 ± 0.14	7.95 ± 0.18	7.61 ± 0.18

Table 6.4: Run IIa predicted and observed number of events in the e+jets channel after 1 tag with NN medium. Errors are statistical only.

Contribution	= 2 Jets	= 3 Jets	≥ 4 Jets
Observed Events	42	45	30
$\sum$ Backgrounds	35.16 ± 1.07	37.73 ± 1.16	35.60 ± 0.91
Semileptonic $t\bar{t}$	4.36 ± 0.12	24.25 ± 0.28	31.00 ± 0.29
Dilepton $t\bar{t}$	4.84 ± 0.05	2.89 ± 0.03	0.94 ± 0.02
Single Top	3.52 ± 0.04	1.94 ± 0.03	0.69 ± 0.02
Multijet	1.59 ± 1.01	2.50 ± 1.11	1.08 ± 0.86
$W + lq$+Jets	0.44 ± 0.00	0.16 ± 0.00	0.04 ± 0.00
$W + bb$+Jets	16.86 ± 0.33	4.26 ± 0.14	1.10 ± 0.06
$W + cc$+Jets	1.93 ± 0.05	0.74 ± 0.03	0.34 ± 0.03
$Z + lq$+Jets	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00
$Z + bb$+Jets	0.54 ± 0.05	0.58 ± 0.04	0.20 ± 0.02
$Z + cc$+Jets	0.03 ± 0.01	0.08 ± 0.01	0.04 ± 0.01
WW	0.11 ± 0.01	0.08 ± 0.01	0.08 ± 0.01
WZ	0.85 ± 0.04	0.20 ± 0.02	0.04 ± 0.01
ZZ	0.03 ± 0.01	0.04 ± 0.01	0.02 ± 0.01
Z'(m=350 GeV)	0.59 ± 0.04	2.70 ± 0.09	3.97 ± 0.11
Z'(m=400 GeV)	0.49 ± 0.03	2.84 ± 0.08	4.54 ± 0.11
Z'(m=450 GeV)	0.42 ± 0.03	2.92 ± 0.08	5.25 ± 0.12
Z'(m=500 GeV)	0.35 ± 0.03	3.12 ± 0.09	5.66 ± 0.13
Z'(m=550 GeV)	0.35 ± 0.02	3.24 ± 0.08	6.04 ± 0.11
Z'(m=600 GeV)	0.41 ± 0.03	3.12 ± 0.09	6.30 ± 0.13
Z'(m=650 GeV)	0.43 ± 0.03	3.41 ± 0.09	6.02 ± 0.13
Z'(m=750 GeV)	0.55 ± 0.04	3.64 ± 0.11	5.37 ± 0.14
Z'(m=850 GeV)	0.78 ± 0.04	3.52 ± 0.09	4.83 ± 0.12
Z'(m=1000 GeV)	0.99 ± 0.04	3.44 ± 0.10	3.92 ± 0.11

Table 6.5: Run IIa predicted and observed number of events in the e+jets channel after ≥ 2 tags with NN medium. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	$\geq$ 4 Jets
Observed Events	14714	6052	1360	385
$\sum$ Backgrounds	14713.68 ± 84.40	6049.71 ± 33.34	1352.99 ± 11.10	377.33 ± 7.27
Semileptonic $t\bar{t}$	2.98 ± 0.22	36.76 ± 0.70	113.80 ± 1.13	124.54 ± 1.07
Dilepton $t\bar{t}$	5.86 ± 0.10	25.29 ± 0.20	12.98 ± 0.13	3.95 ± 0.06
Single Top	12.25 ± 0.14	32.89 ± 0.24	11.68 ± 0.14	2.87 ± 0.08
Multijet	231.33 ± 71.85	142.20 ± 25.72	63.38 ± 8.73	16.81 ± 6.62
$W + lq$+Jets	12377.44 ± 41.53	4411.68 ± 19.16	766.42 ± 5.65	128.16 ± 2.12
$W + bb$+Jets	235.50 ± 2.24	213.13 ± 2.14	61.64 ± 1.04	14.10 ± 0.43
$W + cc$+Jets	540.13 ± 4.51	467.31 ± 4.15	130.36 ± 1.93	32.80 ± 0.88
$Z + lq$+Jets	1080.98 ± 14.29	459.35 ± 7.52	121.09 ± 2.82	35.67 ± 1.42
$Z + bb$+Jets	34.36 ± 0.67	28.33 ± 0.67	9.59 ± 0.38	3.03 ± 0.21
$Z + cc$+Jets	81.02 ± 2.35	58.98 ± 1.81	19.61 ± 0.87	6.66 ± 0.52
WW	84.81 ± 0.61	139.28 ± 0.79	33.69 ± 0.40	6.84 ± 0.18
WZ	14.65 ± 0.31	29.58 ± 0.44	7.23 ± 0.21	1.54 ± 0.10
ZZ	2.36 ± 0.08	4.95 ± 0.11	1.52 ± 0.06	0.36 ± 0.03
Z'(m=350 GeV)	0.18 ± 0.03	3.42 ± 0.17	11.59 ± 0.31	12.44 ± 0.32
Z'(m=400 GeV)	0.21 ± 0.04	3.65 ± 0.16	12.11 ± 0.30	15.11 ± 0.33
Z'(m=450 GeV)	0.23 ± 0.04	3.57 ± 0.17	12.77 ± 0.30	18.11 ± 0.37
Z'(m=500 GeV)	0.23 ± 0.04	3.26 ± 0.17	13.80 ± 0.36	19.83 ± 0.41
Z'(m=550 GeV)	0.25 ± 0.04	3.26 ± 0.14	13.63 ± 0.30	20.91 ± 0.37
Z'(m=600 GeV)	0.27 ± 0.05	3.60 ± 0.16	14.20 ± 0.32	21.64 ± 0.40
Z'(m=650 GeV)	0.25 ± 0.03	3.94 ± 0.16	14.08 ± 0.31	21.29 ± 0.39
Z'(m=750 GeV)	0.42 ± 0.06	4.78 ± 0.21	15.16 ± 0.37	20.81 ± 0.44
Z'(m=850 GeV)	0.43 ± 0.06	5.52 ± 0.20	15.06 ± 0.34	18.51 ± 0.37
Z'(m=1000 GeV)	0.57 ± 0.06	6.54 ± 0.23	13.86 ± 0.32	15.32 ± 0.34

Table 6.6: Run IIa predicted and observed number of events in the μ+jets channel before tagging. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	$\geq$ 4 Jets
Observed Events	347	336	147	106
$\sum$ Backgrounds	331.65 ± 3.93	320.71 ± 4.23	147.33 ± 2.96	78.69 ± 2.44
Semileptonic $t\bar{t}$	0.94 ± 0.08	16.51 ± 0.34	52.68 ± 0.54	56.41 ± 0.49
Dilepton $t\bar{t}$	2.83 ± 0.05	11.83 ± 0.10	6.06 ± 0.06	1.81 ± 0.03
Single Top	5.515 ± 0.07	15.20 ± 0.12	5.41 ± 0.07	1.32 ± 0.04
Multijet	11.10 ± 3.57	15.90 ± 3.98	9.29 ± 2.84	0.00 ± 2.37
$W + lq$+Jets	171.45 ± 1.19	92.27 ± 0.64	20.43 ± 0.21	4.19 ± 0.08
$W + bb$+Jets	74.66 ± 0.93	81.35 ± 0.96	23.97 ± 0.46	5.66 ± 0.19
$W + cc$+Jets	33.75 ± 0.43	51.00 ± 0.60	16.69 ± 0.31	5.08 ± 0.16
$Z + lq$+Jets	9.09 ± 0.14	6.45 ± 0.11	2.45 ± 0.06	0.98 ± 0.04
$Z + bb$+Jets	11.78 ± 0.27	11.02 ± 0.30	4.00 ± 0.17	1.26 ± 0.09
$Z + cc$+Jets	6.00 ± 0.25	6.73 ± 0.29	2.64 ± 0.16	1.05 ± 0.09
WW	3.17 ± 0.04	8.65 ± 0.07	2.63 ± 0.05	0.66 ± 0.03
WZ	1.50 ± 0.08	3.18 ± 0.10	0.87 ± 0.05	0.22 ± 0.03
ZZ	0.24 ± 0.02	0.60 ± 0.03	0.23 ± 0.02	0.05 ± 0.01
Z'(m=350 GeV)	0.08 ± 0.02	1.57 ± 0.08	5.41 ± 0.15	5.75 ± 0.15
Z'(m=400 GeV)	0.08 ± 0.02	1.59 ± 0.08	5.62 ± 0.14	6.88 ± 0.15
Z'(m=450 GeV)	0.06 ± 0.01	1.64 ± 0.08	5.89 ± 0.14	8.25 ± 0.17
Z'(m=500 GeV)	0.08 ± 0.02	1.46 ± 0.08	6.29 ± 0.17	8.94 ± 0.19
Z'(m=550 GeV)	0.07 ± 0.02	1.46 ± 0.06	6.32 ± 0.14	9.48 ± 0.17
Z'(m=600 GeV)	0.07 ± 0.01	1.60 ± 0.08	6.46 ± 0.15	9.86 ± 0.19
Z'(m=650 GeV)	0.05 ± 0.01	1.81 ± 0.08	6.42 ± 0.15	9.68 ± 0.18
Z'(m=750 GeV)	0.12 ± 0.02	2.16 ± 0.10	6.91 ± 0.17	9.49 ± 0.21
Z'(m=850 GeV)	0.15 ± 0.03	2.53 ± 0.10	6.88 ± 0.16	8.41 ± 0.17
Z'(m=1000 GeV)	0.18 ± 0.02	3.07 ± 0.11	6.26 ± 0.15	6.98 ± 0.16

Table 6.7: Run IIa predicted and observed number of events in the μ+jets channel after 1 tag with NN medium. Errors are statistical only.

Contribution	= 2 Jets	= 3 Jets	≥ 4 Jets
Observed Events	30	38	35
$\sum$ Backgrounds	31.40 ± 1.02	33.17 ± 1.19	34.15 ± 0.93
Semileptonic $t\bar{t}$	2.89 ± 0.10	20.81 ± 0.27	30.48 ± 0.30
Dilepton $t\bar{t}$	5.04 ± 0.05	2.83 ± 0.03	0.88 ± 0.02
Single Top	2.82 ± 0.03	1.64 ± 0.03	0.53 ± 0.01
Multijet	0.47 ± 0.96	0.61 ± 1.15	0.00 ± 0.88
$W + lq$+Jets	0.43 ± 0.00	0.18 ± 0.00	0.05 ± 0.00
$W + bb$+Jets	15.16 ± 0.31	4.91 ± 0.17	1.23 ± 0.07
$W + cc$+Jets	1.80 ± 0.04	0.81 ± 0.03	0.38 ± 0.02
$Z + lq$+Jets	0.02 ± 0.00	0.02 ± 0.00	0.01 ± 0.00
$Z + bb$+Jets	1.58 ± 0.07	0.85 ± 0.06	0.38 ± 0.04
$Z + cc$+Jets	0.25 ± 0.03	0.15 ± 0.02	0.07 ± 0.01
WW	0.07 ± 0.00	0.09 ± 0.01	0.06 ± 0.01
WZ	0.69 ± 0.04	0.22 ± 0.02	0.07 ± 0.01
ZZ	0.14 ± 0.01	0.06 ± 0.01	0.01 ± 0.00
Z'(m=350 GeV)	0.28 ± 0.02	2.08 ± 0.07	2.93 ± 0.09
Z'(m=400 GeV)	0.29 ± 0.02	2.11 ± 0.07	3.51 ± 0.09
Z'(m=450 GeV)	0.28 ± 0.02	2.16 ± 0.07	4.46 ± 0.11
Z'(m=500 GeV)	0.25 ± 0.03	2.47 ± 0.08	4.80 ± 0.12
Z'(m=550 GeV)	0.26 ± 0.02	2.56 ± 0.07	5.15 ± 0.11
Z'(m=600 GeV)	0.27 ± 0.02	2.77 ± 0.08	5.53 ± 0.12
Z'(m=650 GeV)	0.32 ± 0.02	2.82 ± 0.08	5.42 ± 0.11
Z'(m=750 GeV)	0.44 ± 0.03	3.24 ± 0.10	5.28 ± 0.13
Z'(m=850 GeV)	0.64 ± 0.03	3.35 ± 0.09	4.56 ± 0.11
Z'(m=1000 GeV)	0.95 ± 0.05	3.16 ± 0.09	3.88 ± 0.10

Table 6.8: Run IIa predicted and observed number of events in the μ+jets channel after ≥ 2 tags with NN medium. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	$\geq$ 4 Jets
Observed Events	39877	15962	3692	986
$\sum$ Backgrounds	39877.86 ± 83.20	15961.93 ± 41.10	3692.00 ± 17.55	986.00 ± 7.12
Semileptonic $t\bar{t}$	9.32 ± 0.25	101.91 ± 0.82	324.63 ± 1.46	411.74 ± 1.59
Dilepton $t\bar{t}$	34.19 ± 0.24	133.72 ± 0.47	78.07 ± 0.35	22.75 ± 0.17
Single Top	42.08 ± 0.35	96.27 ± 0.53	30.76 ± 0.31	7.99 ± 0.17
Multijet	1957.66 ± 25.16	1503.44 ± 18.85	478.56 ± 10.13	127.36 ± 5.22
$W + lq$+Jets	32802.04 ± 74.61	9973.50 ± 28.31	1681.47 ± 10.16	207.67 ± 2.53
$W + bb$+Jets	1008.08 ± 7.67	771.11 ± 6.26	198.51 ± 2.63	30.81 ± 0.78
$W + cc$+Jets	2321.39 ± 19.32	1637.83 ± 14.91	428.39 ± 6.43	67.07 ± 1.93
$Z + lq$+Jets	1171.06 ± 16.41	1014.79 ± 15.51	254.77 ± 6.57	55.75 ± 2.81
$Z + bb$+Jets	88.20 ± 1.69	118.76 ± 2.01	51.49 ± 1.34	15.51 ± 0.75
$Z + cc$+Jets	88.20 ± 1.69	118.76 ± 2.01	51.49 ± 1.34	15.51 ± 0.75
WW	306.22 ± 3.79	412.45 ± 4.43	92.66 ± 2.12	19.07 ± 0.99
WZ	44.63 ± 0.84	74.33 ± 1.10	18.22 ± 0.55	3.92 ± 0.27
ZZ	4.79 ± 0.18	5.05 ± 0.19	2.96 ± 0.15	0.85 ± 0.08
Z'(m=350 GeV)	6.61 ± 0.44	29.31 ± 0.95	46.79 ± 1.20	47.13 ± 1.24
Z'(m=400 GeV)	5.98 ± 0.35	28.53 ± 0.75	51.54 ± 1.02	56.59 ± 1.08
Z'(m=450 GeV)	5.22 ± 0.39	29.02 ± 0.93	52.78 ± 1.27	62.79 ± 1.39
Z'(m=500 GeV)	5.42 ± 0.31	29.60 ± 0.72	54.28 ± 0.98	66.50 ± 1.09
Z'(m=550 GeV)	5.42 ± 0.40	29.45 ± 0.94	54.74 ± 1.28	69.30 ± 1.47
Z'(m=600 GeV)	5.74 ± 0.38	28.66 ± 0.81	54.86 ± 1.12	71.45 ± 1.32
Z'(m=650 GeV)	6.37 ± 0.41	31.98 ± 0.94	55.80 ± 1.27	69.54 ± 1.43
Z'(m=750 GeV)	8.06 ± 0.51	37.08 ± 1.07	57.26 ± 1.34	63.55 ± 1.42
Z'(m=850 GeV)	9.31 ± 0.42	38.79 ± 0.87	56.39 ± 1.06	54.84 ± 1.05
Z'(m=1000 GeV)	11.82 ± 0.61	42.15 ± 1.15	53.84 ± 1.31	49.70 ± 1.26

Table 6.9: Run IIb predicted and observed number of events in the e+jets channel before tagging. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	≥ 4 Jets
Observed Events	1241	997	470	244
$\sum$ Backgrounds	1320.28 ± 6.42	1212.55 ± 6.42	504.66 ± 3.65	267.78 ± 2.17
Semileptonic $t\bar{t}$	2.80 ± 0.09	44.94 ± 0.38	152.18 ± 0.70	191.75 ± 0.75
Dilepton $t\bar{t}$	15.44 ± 0.11	63.74 ± 0.23	37.36 ± 0.17	10.83 ± 0.08
Single Top	17.31 ± 0.16	44.53 ± 0.26	14.30 ± 0.15	3.74 ± 0.08
Multijet	72.08 ± 4.37	104.39 ± 4.81	35.78 ± 3.02	11.95 ± 1.92
$W + lq$+Jets	665.31 ± 2.54	320.40 ± 1.24	70.40 ± 0.50	10.92 ± 0.14
$W + bb$+Jets	323.45 ± 3.27	306.76 ± 2.91	81.93 ± 1.22	13.03 ± 0.35
$W + cc$+Jets	168.48 ± 2.12	214.82 ± 2.64	67.56 ± 1.26	12.31 ± 0.40
$Z + lq$+Jets	17.59 ± 0.25	25.47 ± 0.39	9.02 ± 0.23	2.54 ± 0.12
$Z + bb$+Jets	9.53 ± 0.29	22.36 ± 0.46	12.08 ± 0.34	3.85 ± 0.21
$Z + cc$+Jets	9.53 ± 0.29	22.36 ± 0.46	12.08 ± 0.34	3.85 ± 0.21
WW	13.51 ± 0.27	31.97 ± 0.46	8.85 ± 0.27	2.25 ± 0.15
WZ	4.92 ± 0.18	10.13 ± 0.25	2.62 ± 0.12	0.59 ± 0.06
ZZ	0.32 ± 0.03	0.67 ± 0.04	0.51 ± 0.04	0.16 ± 0.02
Z'(m=350 GeV)	2.84 ± 0.21	13.66 ± 0.46	22.00 ± 0.58	21.89 ± 0.58
Z'(m=400 GeV)	2.59 ± 0.16	13.07 ± 0.36	24.31 ± 0.49	26.30 ± 0.51
Z'(m=450 GeV)	2.19 ± 0.18	13.11 ± 0.44	24.60 ± 0.61	29.08 ± 0.65
Z'(m=500 GeV)	2.20 ± 0.14	13.28 ± 0.34	25.12 ± 0.46	31.03 ± 0.52
Z'(m=550 GeV)	2.06 ± 0.17	13.15 ± 0.44	25.69 ± 0.61	32.21 ± 0.69
Z'(m=600 GeV)	1.99 ± 0.15	12.92 ± 0.38	25.31 ± 0.53	33.09 ± 0.62
Z'(m=650 GeV)	2.21 ± 0.16	14.07 ± 0.43	25.79 ± 0.60	32.28 ± 0.67
Z'(m=750 GeV)	2.63 ± 0.19	16.01 ± 0.48	26.29 ± 0.63	29.15 ± 0.67
Z'(m=850 GeV)	2.56 ± 0.13	16.21 ± 0.39	25.60 ± 0.50	25.09 ± 0.49
Z'(m=1000 GeV)	2.90 ± 0.18	16.84 ± 0.49	23.82 ± 0.60	22.54 ± 0.58

Table 6.10: Run IIb predicted and observed number of events in the e+jets channel after 1 tag with NN medium. Errors are statistical only.

Contribution	= 2 Jets	= 3 Jets	≥ 4 Jets
Observed Events	107	109	88
$\sum$ Backgrounds	116.26 ± 1.44	105.03 ± 1.27	113.05 ± 1.04
Semileptonic $t\bar{t}$	7.29 ± 0.10	54.86 ± 0.31	99.06 ± 0.43
Dilepton $t\bar{t}$	23.34 ± 0.10	16.05 ± 0.08	5.05 ± 0.04
Single Top	7.82 ± 0.06	4.04 ± 0.05	1.43 ± 0.04
Multijet	1.53 ± 0.99	2.76 ± 1.12	1.24 ± 0.93
$W + lq$+Jets	2.45 ± 0.01	1.00 ± 0.01	0.23 ± 0.00
$W + bb$+Jets	58.26 ± 1.00	17.78 ± 0.46	3.14 ± 0.13
$W + cc$+Jets	9.18 ± 0.20	4.18 ± 0.13	1.10 ± 0.06
$Z + lq$+Jets	0.17 ± 0.00	0.12 ± 0.00	0.05 ± 0.00
$Z + bb$+Jets	1.68 ± 0.09	1.60 ± 0.08	0.70 ± 0.06
$Z + cc$+Jets	1.68 ± 0.09	1.60 ± 0.08	0.70 ± 0.06
WW	0.41 ± 0.02	0.32 ± 0.04	0.16 ± 0.03
WZ	2.34 ± 0.09	0.60 ± 0.05	0.13 ± 0.02
ZZ	0.11 ± 0.01	0.13 ± 0.02	0.04 ± 0.01
Z'(m=350 GeV)	3.60 ± 0.17	8.13 ± 0.26	11.06 ± 0.33
Z'(m=400 GeV)	3.66 ± 0.14	8.95 ± 0.22	13.25 ± 0.29
Z'(m=450 GeV)	3.47 ± 0.16	8.99 ± 0.27	14.71 ± 0.37
Z'(m=500 GeV)	3.50 ± 0.12	8.95 ± 0.20	15.49 ± 0.29
Z'(m=550 GeV)	3.40 ± 0.15	9.24 ± 0.27	15.86 ± 0.38
Z'(m=600 GeV)	3.04 ± 0.12	8.92 ± 0.22	16.32 ± 0.34
Z'(m=650 GeV)	3.14 ± 0.13	9.12 ± 0.25	15.57 ± 0.36
Z'(m=750 GeV)	3.24 ± 0.14	9.00 ± 0.26	13.55 ± 0.34
Z'(m=850 GeV)	3.04 ± 0.10	8.50 ± 0.20	11.27 ± 0.25
Z'(m=1000 GeV)	2.78 ± 0.12	7.11 ± 0.22	9.55 ± 0.28

Table 6.11: Run IIb predicted and observed number of events in the e+jets channel after ≥ 2 tags with NN medium. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	$\geq$ 4 Jets
Observed Events	27265	12015	2829	826
$\sum$ Backgrounds	27266.14 ± 58.36	12015.14 ± 29.57	2829.00 ± 12.61	826.00 ± 5.22
Semileptonic $t\bar{t}$	3.45 ± 0.13	59.21 ± 0.56	224.99 ± 1.11	311.10 ± 1.26
Dilepton $t\bar{t}$	19.02 ± 0.15	94.79 ± 0.35	58.71 ± 0.27	16.48 ± 0.13
Single Top	24.07 ± 0.23	64.65 ± 0.38	21.15 ± 0.23	5.29 ± 0.12
Multijet	66.25 ± 8.18	57.50 ± 6.03	14.82 ± 3.18	6.14 ± 1.81
$Z + lq$+Jets	1999.77 ± 18.31	819.10 ± 9.60	185.80 ± 4.11	30.01 ± 1.52
$Z + bb$+Jets	152.26 ± 2.01	140.48 ± 2.00	43.73 ± 1.11	9.35 ± 0.49
$Z + cc$+Jets	152.26 ± 2.01	140.48 ± 2.00	43.73 ± 1.11	9.35 ± 0.49
$W + lq$+Jets	22445.09 ± 52.80	8240.90 ± 23.14	1560.03 ± 9.20	278.65 ± 3.15
$W + bb$+Jets	647.10 ± 5.19	633.67 ± 5.19	179.40 ± 2.37	43.75 ± 1.05
$W + cc$+Jets	1499.38 ± 13.17	1362.22 ± 12.70	402.91 ± 5.90	97.25 ± 2.81
WW	218.09 ± 2.80	324.35 ± 3.52	74.20 ± 1.72	14.95 ± 0.79
WZ	34.40 ± 0.65	68.31 ± 0.95	16.35 ± 0.48	3.04 ± 0.21
ZZ	5.00 ± 0.17	9.46 ± 0.24	3.16 ± 0.14	0.65 ± 0.06
Z'(m=350 GeV)	3.28 ± 0.28	18.83 ± 0.69	32.71 ± 0.90	31.66 ± 0.91
Z'(m=400 GeV)	2.94 ± 0.20	18.70 ± 0.55	35.76 ± 0.77	39.23 ± 0.82
Z'(m=450 GeV)	2.73 ± 0.25	19.79 ± 0.70	38.67 ± 0.98	47.04 ± 1.10
Z'(m=500 GeV)	2.97 ± 0.19	20.33 ± 0.54	40.32 ± 0.77	50.79 ± 0.87
Z'(m=550 GeV)	3.11 ± 0.27	20.02 ± 0.70	41.41 ± 1.00	51.39 ± 1.14
Z'(m=600 GeV)	2.84 ± 0.23	21.34 ± 0.65	41.83 ± 0.90	54.83 ± 1.05
Z'(m=650 GeV)	2.53 ± 0.23	20.99 ± 0.70	39.56 ± 0.96	57.05 ± 1.18
Z'(m=750 GeV)	2.90 ± 0.27	21.52 ± 0.74	41.11 ± 1.02	55.35 ± 1.21
Z'(m=850 GeV)	3.30 ± 0.22	23.68 ± 0.61	45.40 ± 0.87	48.13 ± 0.90
Z'(m=1000 GeV)	2.77 ± 0.24	23.89 ± 0.77	41.07 ± 1.03	44.70 ± 1.11

Table 6.12: Run IIb predicted and observed number of events in the μ+jets channel before tagging. Errors are statistical only.

Contribution	= 1 Jets	= 2 Jets	= 3 Jets	≥ 4 Jets
Observed Events	891	759	338	217
$\sum$ Backgrounds	857.28 ± 3.45	903.24 ± 3.86	381.21 ± 2.11	213.78 ± 1.32
Semileptonic $t\bar{t}$	1.00 ± 0.05	25.95 ± 0.26	104.38 ± 0.53	143.39 ± 0.59
Dilepton $t\bar{t}$	8.41 ± 0.07	44.59 ± 0.17	27.77 ± 0.13	7.73 ± 0.06
Single Top	9.74 ± 0.10	29.78 ± 0.19	9.77 ± 0.11	2.46 ± 0.06
Multijet	1.29 ± 1.30	5.69 ± 1.53	1.59 ± 1.08	0.00 ± 0.83
$W + lq$+Jets	440.14 ± 1.72	258.70 ± 0.96	64.62 ± 0.45	14.55 ± 0.18
$W + bb$+Jets	201.54 ± 2.18	248.84 ± 2.40	73.74 ± 1.10	18.73 ± 0.49
$W + cc$+Jets	108.22 ± 1.43	175.19 ± 2.24	62.24 ± 1.14	17.87 ± 0.59
$Z + lq$+Jets	28.38 ± 0.28	19.91 ± 0.25	6.35 ± 0.15	1.39 ± 0.08
$Z + bb$+Jets	22.19 ± 0.39	29.59 ± 0.49	10.51 ± 0.30	2.57 ± 0.15
$Z + cc$+Jets	22.19 ± 0.39	29.59 ± 0.49	10.51 ± 0.30	2.57 ± 0.15
WW	9.78 ± 0.20	25.28 ± 0.37	7.01 ± 0.22	1.93 ± 0.14
WZ	3.85 ± 0.14	8.76 ± 0.20	2.25 ± 0.10	0.47 ± 0.05
ZZ	0.57 ± 0.04	1.37 ± 0.06	0.47 ± 0.03	0.12 ± 0.02
Z'(m=350 GeV)	1.36 ± 0.13	8.57 ± 0.33	15.08 ± 0.43	14.70 ± 0.43
Z'(m=400 GeV)	1.17 ± 0.09	8.45 ± 0.26	16.58 ± 0.37	18.05 ± 0.38
Z'(m=450 GeV)	1.14 ± 0.11	8.92 ± 0.33	18.03 ± 0.47	21.73 ± 0.52
Z'(m=500 GeV)	1.19 ± 0.09	9.20 ± 0.25	18.62 ± 0.37	23.37 ± 0.41
Z'(m=550 GeV)	1.26 ± 0.12	8.93 ± 0.33	19.00 ± 0.47	23.62 ± 0.53
Z'(m=600 GeV)	1.11 ± 0.10	9.63 ± 0.31	19.28 ± 0.43	25.02 ± 0.49
Z'(m=650 GeV)	0.94 ± 0.10	9.29 ± 0.33	18.21 ± 0.46	26.27 ± 0.55
Z'(m=750 GeV)	1.01 ± 0.11	9.58 ± 0.34	18.86 ± 0.48	25.37 ± 0.57
Z'(m=850 GeV)	1.06 ± 0.08	10.34 ± 0.28	20.58 ± 0.41	21.88 ± 0.42
Z'(m=1000 GeV)	0.84 ± 0.09	10.02 ± 0.34	18.54 ± 0.48	20.19 ± 0.51

Table 6.13: Run IIb predicted and observed number of events in the μ+jets channel after 1 tag with NN medium. Errors are statistical only.

Contribution	= 2 Jets	= 3 Jets	≥ 4 Jets
Observed Events	70	72	99
$\sum$ Backgrounds	93.38 ± 0.96	77.96 ± 0.66	87.18 ± 0.61
Semileptonic $t\bar{t}$	4.11 ± 0.07	38.25 ± 0.24	74.67 ± 0.34
Dilepton $t\bar{t}$	16.74 ± 0.08	12.08 ± 0.06	3.60 ± 0.03
Single Top	5.98 ± 0.05	3.00 ± 0.04	0.97 ± 0.03
Multijet	0.00 ± 0.38	0.00 ± 0.42	0.00 ± 0.46
$W + lq$+Jets	2.01 ± 0.01	0.92 ± 0.01	0.31 ± 0.00
$W + bb$+Jets	47.34 ± 0.84	15.79 ± 0.41	4.63 ± 0.19
$W + cc$+Jets	7.29 ± 0.16	3.81 ± 0.12	1.69 ± 0.09
$Z + lq$+Jets	0.13 ± 0.00	0.08 ± 0.00	0.03 ± 0.00
$Z + bb$+Jets	3.55 ± 0.12	1.58 ± 0.08	0.49 ± 0.04
$Z + cc$+Jets	3.55 ± 0.12	1.58 ± 0.08	0.49 ± 0.04
WW	0.38 ± 0.02	0.26 ± 0.02	0.17 ± 0.03
WZ	1.97 ± 0.08	0.47 ± 0.04	0.11 ± 0.02
ZZ	0.33 ± 0.02	0.12 ± 0.01	0.03 ± 0.01
Z'(m=350 GeV)	2.52 ± 0.13	5.71 ± 0.20	7.58 ± 0.25
Z'(m=400 GeV)	2.52 ± 0.10	6.22 ± 0.17	9.17 ± 0.22
Z'(m=450 GeV)	2.47 ± 0.12	6.73 ± 0.21	11.16 ± 0.29
Z'(m=500 GeV)	2.74 ± 0.10	6.95 ± 0.17	11.92 ± 0.23
Z'(m=550 GeV)	2.48 ± 0.12	7.03 ± 0.21	11.90 ± 0.30
Z'(m=600 GeV)	2.70 ± 0.11	7.22 ± 0.19	12.62 ± 0.28
Z'(m=650 GeV)	2.33 ± 0.11	6.87 ± 0.20	13.19 ± 0.31
Z'(m=750 GeV)	2.36 ± 0.11	7.00 ± 0.21	12.33 ± 0.31
Z'(m=850 GeV)	2.40 ± 0.08	7.29 ± 0.17	10.35 ± 0.22
Z'(m=1000 GeV)	2.09 ± 0.10	6.25 ± 0.19	8.97 ± 0.26

Table 6.14: Run IIb predicted and observed number of events in the μ+jets channel after ≥ 2 tags with NN medium. Errors are statistical only.

6.4 Invariant Mass Reconstruction

Various methods have been evaluated to compute the $t\bar{t}$ invariant mass $M_{t\bar{t}}$ from the reconstructed physic objects. In all cases the z component of the neutrino momentum has to be inferred using constraints, as it cannot be directly measured in experiments using hadron colliders.

A kinematic fit to the full event based on `HitFit` [176] has been investigated. Only exactly four jets may be passed to this fit. Thus at least four jets are required in this method, in case of more than four jets the four leading jets in p_T are used.

The most direct method is to use the measured $\not\!\!E_{\mathrm{T}}$ and lepton momentum. Requiring, that the neutrino and lepton stem from the same W boson leads to a quadratic equation, that can be solved for the z-component of the neutrino momentum. This method is possible for almost any number of jets.

The various methods are explained in detail and are compared in the following sections.

6.4.1 Constraint Fit

The constraint fit method has been used in previous Run I [177] and Run II [176] analyses. It uses the four vectors of the $t\bar{t}$ decay products: The four (quark) jets, the lepton and the neutrino. The four vectors are combined by using the SQUAW algorithm [178]. The kinematic fit is implemented in a DØ tool called `HitFit`. The exact treatment can be found in [179].

The program `HitFit` matches a number of constraints by changing the vector of the measured four vectors $\vec{x}^m$ of the objects as little as possible. The used constraints are:

- Lepton and $\not\!\!E_{\mathrm{T}}$ must form the invariant mass of the W boson.
- Two jets form the invariant mass of the W boson.
- The two reconstructed top quark masses have to be equal.

The top quark mass could be fixed, but since we vary the top quark mass for systematic uncertainties, every sample would have to be calculated with multiple top quark masses and thus fixing the mass gives no significant improvement in the fit.

The program uses the input vectors given and tries to match the constraints above by changing the vectors $\vec{x}$ to a set of modified vectors $\vec{x}_{\mathrm{m}}$, where the change

of the original vectors is tried to be minimized by a minimization of a χ^2 given by:

$$\chi^2 = (\vec{x} - \vec{x}^m)^T \mathbf{G} (\vec{x} - \vec{x}^m), \tag{6.6}$$

where $\mathbf{G}$ is the inverse error matrix, which is diagonal since the input vectors are transformed such that their errors are uncorrelated.

All objects, but the neutrino can be fully reconstructed in the detector. The neutrino momentum can only be calculated from $\vec{\not{E}}$, which lacks the z component. To solve this, one uses the assumption that both top quarks have the same mass, which leads to:

$$\begin{aligned} 0 &= [(p_z^c)^2 - (E^c)^2](p_z^\nu)^2 + \alpha p_z^c p_z^\nu - (E^c p_T^\nu)^2 + \frac{\alpha^2}{4} \qquad \text{with} \\ \alpha &= m_t^2 - m_c^2 + 2\vec{p}_T^{\,\nu} \cdot \vec{p}_T^{\,c}, \end{aligned} \tag{6.7}$$

where c is the object composed from the lepton and the b-jet. This leads to two solutions for the longitudinal neutrino momentum p_z^ν. Both values are used in the fit and the one resulting in the lower χ^2 solution is kept.

As the parton assignment to the jets is not known, 4!=24 solutions for the assignment are possible. As both light jets do not need to be distinguished, the number can be reduced to 12, without loosing any information. In addition, b-tags can be used to constrain the combinatorics to six in one b-tag events or two combinatorial possible solutions in two b-tag events.

While the percentage of events in which the kinematical fits converge is between 80 %–100 %, the correct solution is achieved in 37.5 % to 66 % of the cases. The fit converges if the χ^2 is below ten. In general hogher resonance masses yield lower convergence of the fit and lower correct assignments. Though these numbers seemslow, they are much higher than a random choice, which would yield only 2 %. Non converging solutions get removed from the selection. Several methods to exploit the b-tag information of the jets are tested, only marginally changing the result.

The mass resolution for the $t\bar{t}$ background is measured to be $\sim$ 43 GeV and similar to the resolution of a Z' boson with a mass of 350 GeV. The width of the reconstructed mass broadens heavily with rising resonance mass.

6.4.2 Direct Approach

In the most direct method the invariant mass is reconstructed by using the 4-vectors of up to four leading jets, the lepton, and the neutrino of the event. The momenta of all objects but the neutrino are measured by the calorimeter. The x- and y-component of the neutrino are given by the $\vec{\not{E}}$ and the z-component

can be calculated by requiring the lepton and the neutrino vectors adding up to give the resulting four vector the invariant mass of a W boson. This leads to a quadratic equation. In case of more than one solution, the one with the lower z-momentum is used which is the correct one in about 70% of all cases, as determined by MC-studies. In case there is no solution, the z-momentum of the neutrino is recalculated by changing the $\not{E}_{\mathrm{T}}$ to force the invariant mass of the lepton and the neutrino to give the W-boson mass. This is appropriate as the $\not{E}_{\mathrm{T}}$ is the variable which has the worst resolution. The amount of forced $\not{E}_{\mathrm{T}}$ recalculations is of the order of 15 % depending on the sample. The $\not{E}_{\mathrm{T}}$ is only recalculated for the invariant mass calculcation. All other distributions show the original $\not{E}_{\mathrm{T}}$.

To compute the invariant $t\bar{t}$ mass, the neutrino and lepton momenta need to be added with the jet momenta. This procedure is applicable for events with any number of jets. Up to four leading jets are used in the computation. If an event features more jets, the four leading jets are used to compute the invariant $t\bar{t}$ mass. The data is separated into events with exactly three jets and events with four or more jets. In case of three jets, in most cases the two light quark jets from the W-boson decay have been reconstructed in a single jet.

The reconstructed invariant $t\bar{t}$ mass of the SM $t\bar{t}$ process as well as for three Z' resonance masses are shown in figure 6.1 and for pure vector- or axial-vector-couplings in figure 6.2. No significant difference can be observed in the $t\bar{t}$ invariant mass spectra of the three resonance types. The reconstructed width of the resonances broadens significantly with higher resonance masses, making it harder to distinguish from the SM background. Higher resonance masses also shift the contributions to a mass range, where one would expect no standard model $t\bar{t}$ background. The plots have been normalized to identical area. The measured contributions for the theoretical models can be seen in chapter 7.

6.4.3 Comparison

Former analyses have used the constraint kinematic fit to improve the results. Using this method has, however, proven to be less efficient for this analysis for various reasons. Only exactly four jets may be passed to this fit. Thus at least four jets are required in this method. Therefore this method is limited to the 4-jet bin, while the direct method is applicable for any number of jet bins. It cannot be used in the 3-jet bin, which is an addition to the last thesis published [180]. The 3-jet bin contains about 50 % $t\bar{t}$ signal after requiring at least one b-tag. This gives an additional amount of signal events on the measurement, especially as the amount of $t\bar{t}$ events in the three jet bin roughly matches the number in the 4-jet bin after b tagging. This leads to almost the double amount of data available

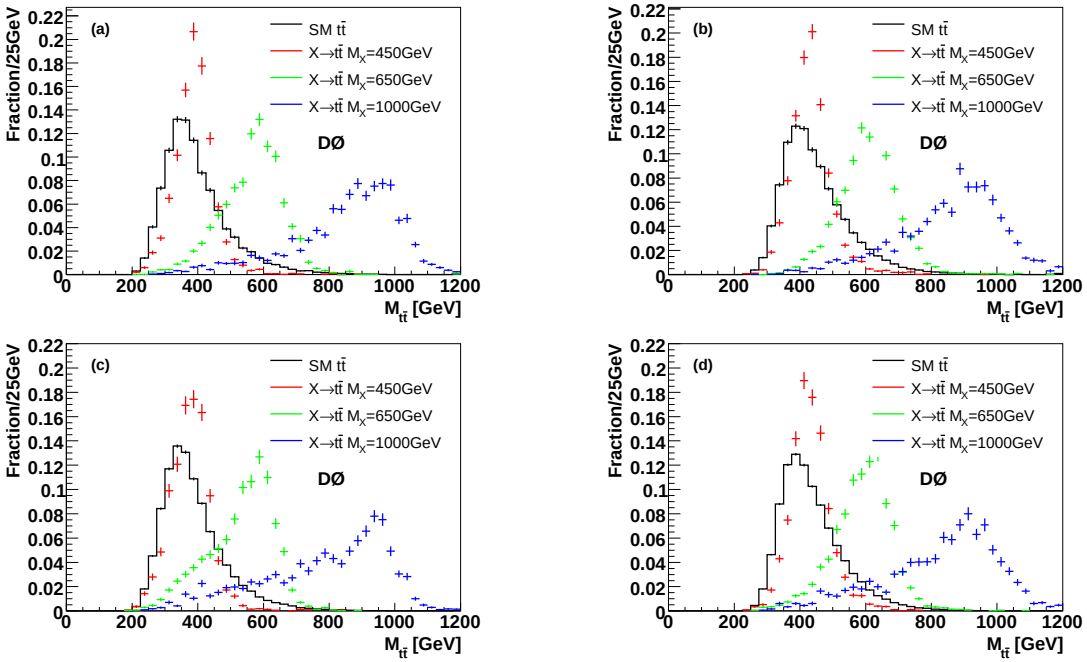


Figure 6.1: Comparison of the shape of the reconstructed invariant $t\bar{t}$ mass between SM and resonant production. The shapes correspond to the combined lepton+jets channel with at least one b-tag. First row corresponds to Run IIa, second to Run IIb. Left column shows the direct reconstruction 3-jets, the right one with 4-jets. Within each plot resonance masses of 450 GeV(red), 650 GeV (green), and 1 TeV(blue) are shown. The curves are area normalized.

compared to previous analyses.

Moreover, in case of higher resonance masses, two or more partons are more likely ending up in one jet, which migrates the events to a lower jet bin. Especially for the 4-jet bin this gives three (or less) jets from the hard interaction and one (or more) additional jets from for example pile-up, underlying event, or ISR. The fit then tries to find a solution, which treats the additional jet as being from the original hard interaction, thus leading to no solution in the fit. As the number of merged partons increases with higher resonance mass, the amount of true 4-jet events decreases in the sample. This effect is demonstrated in [156] where the mass resolution using `HitFit` is seen to deteriorate at higher resonance masses.

Tests have shown, that while the constraint fit with `HitFit` leads to a minor improvement in the mass range of 350 to 400 GeV the results worsen in the region of high invariant mass of 650 GeV or higher. The performance of the different reconstruction methods can also be judged by comparing the expected limits obtained by the reconstructed $M_{t\bar{t}}$ distribution, that are shown in figure 6.3. As

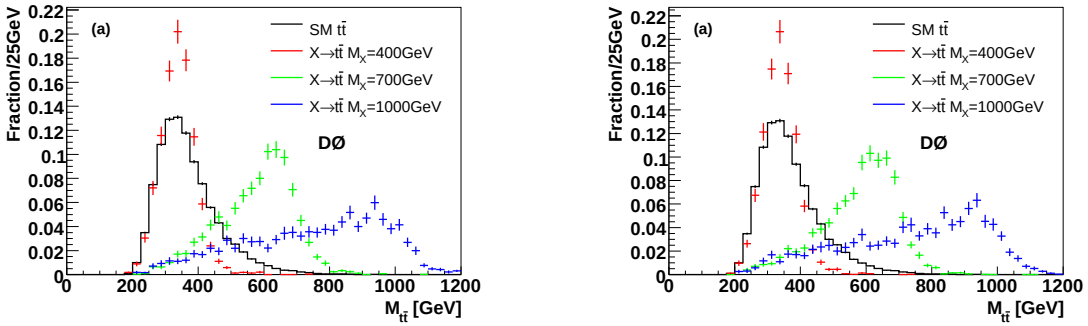


Figure 6.2: Comparison of the shape of the reconstructed invariant $t\bar{t}$ mass between SM and pure vector coupling (left) or axial vector coupling (right) resonant production. The shapes correspond to the combined lepton+jets channel with at least one b-tag. Direct reconstruction 3-jets and within each plot resonance masses of 450 GeV(red), 650 GeV (green), and 1 TeV(blue) are shown. The curves are area normalized.

the main goal of this thesis is to search for $t\bar{t}$ resonances, the method is tuned to have maximal sensitivity for a heavy resonance, as the low mass region can be excluded with the direct approach as well. Therefore the approach with the direct reconstruction is chosen. The reconstruction of the four leading jets is slightly better, than the one with more than four jets, which usually have a fifth jet which is not from the hard interaction. As those are normally low p_{T} jets, the change is minimal.

6.5 Mass Resolution

The mass resolution for the detector is studied in Monte Carlo simulated events by comparing the reconstructed invariant $t\bar{t}$ mass and the information of the partons directly from Monte Carlo. The resolution is computed by taking the difference between the reconstructed mass and the true mass. The RMS of the distribution is then the resolution of the detector.

Table 6.15 shows the mass resolution of the different Z' boson samples. The table shows the width of the generated Z' boson and the mass resolution in GeV and per cent of the mass. The mass resolution is in the order of about $10.7 \pm 2\,\%$ and therefore well above the generated width. The mass resolution for a $t\bar{t}$ sample is about 43 GeV. The width of the Z' boson is therefore undetectable by the detector and can be neglected as systematic effect.

For comparison purposes, the Monte Carlo true information, the reconstructed information and the difference for various Z' boson masses is shown in figure 6.4.

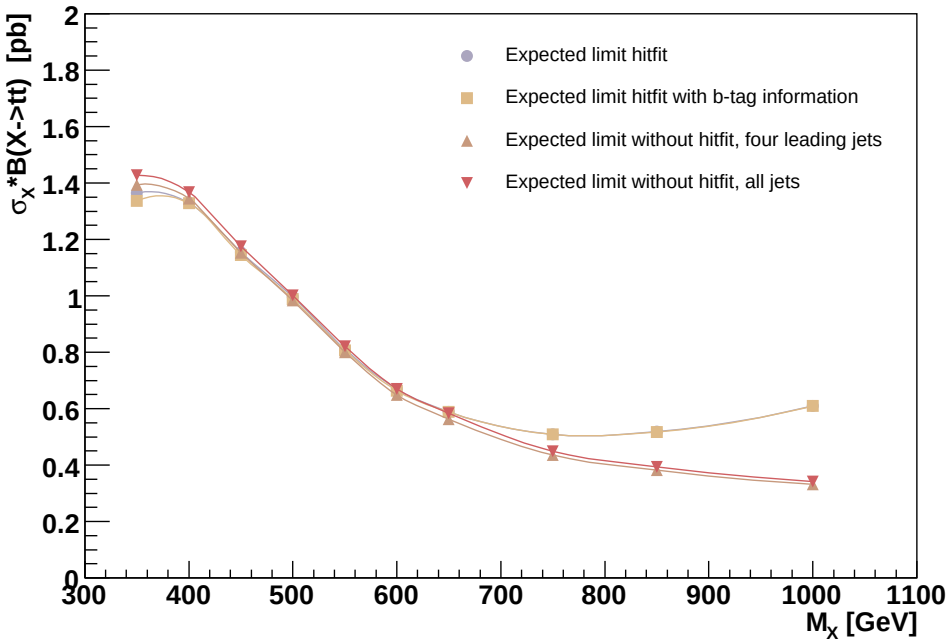


Figure 6.3: Limits obtained with different reconstruction methods for the $t\bar{t}$ invariant mass spectrum. The constraint fits with `HitFit` with and without exploiting the b-tag information, as well as the direct reconstruction for the four leading jets and all jets in the event. The direct reconstuction method though not the best at the low mass region is giving the best sensitivity in the high mass region.

With increasing mass, the radiative tail increases, while the high mass tail decreases due to the lower probability for high x in the PDFs. The generic form of the resonance is still visible in the reconstruction though smeared out. While the mass resolution broadens, the relative mass resolution stays nearly constant.

6.6 Systematic Uncertainty Studies

The result in a physical measurement is not only the central or mean value but also its uncertainty. The estimation of the uncertainty is usually given for the statistical and the systematic uncertainty. Usually, the convolution of all uncertainties including their correlations needs to be taken into account.

Depending on the mass of the studied resonance, the number of events in that range is more or less limited. As can be seen in the next chapter, the analysis is systematically limited in the low mass region, while the high mass region is

Mass [GeV]	Width [GeV]	Mass Resolution [GeV]	Mass Resolution %
350	4,2	42,0	12,0
400	4,8	43,3	10,8
450	5,4	50,3	11,2
500	6,0	52,9	10,6
550	6,6	55,2	10,0
600	7,2	59,7	10,0
650	7,8	64,6	9,9
700	8,4	78,1	11,2
750	9,0	65,2	8,7
850	10,2	107,8	12,7
1000	12,0	124,4	12,4

Table 6.15: Mass resolution of the Z' bosons given in GeV and in % of the rest mass. The Z' boson width of 1.2 % of the rest mass is in all cases smaller than the detector mass resolution.

dominated by the statistical uncertainty as can be seen in figure 7.2.

The following section focuses on the judgment of the systematic uncertainties and their size.

6.6.1 Sources of Systematic Uncertainties

This analysis relies on the shapes and overall normalization of the invariant $M_{t\bar{t}}$ distributions. The systematic uncertainties, which only affect the normalization, include the uncertainties on the MC-to-data correction factors, the cross sections, and the luminosity. Shape and normalization changing systematic uncertainties are the jet energy scale, jet energy resolution, and those related to b-tagging, among others. The following systematic effects have been evaluated on all samples including the signal.

Jet Energy Scale The uncertainty on the jet energy scale is determined by changing the JES correction up and down by one standard deviation [181]. This includes the systematic and statistical shift added in quadrature. The impact on the selection efficiency is taken into account.

Jet Energy Resolution The uncertainty on the jet energy resolution is determined by changing the JER correction up and down by one standard deviation [181].

The impact on the selection efficiency is taken into account.

Jet ID The uncertainty on the jet reconstruction and jet identification efficiency (JetID) is determined by changing the JetID down by one standard deviation [181]. The impact on the preselection efficiency is taken into account. The upward correction is obtained by using an opposite shift of the same size in each bin of the final $M_{t\bar{t}}$ distributions.

$t\bar{t}$ Cross Section The uncertainty on $\sigma_{t\bar{t}}(m_t = 172.4\,\text{GeV})$ which is set to be $^{+0.48}_{-0.67}$ pb, is added in quadrature to the uncertainty $^{+0.28}_{-0.27}$ due to the change of $\sigma_{t\bar{t}}(m_t)$ when varying m_t by $\pm 1.2\,\text{GeV}$ [182]. Thus the total cross section uncertainty is $^{+0.56}_{-0.72}$ pb.

Top Quark Mass Kinematics The uncertainty due to the top quark mass uncertainty is estimated by modifying m_t in the simulation of the dominating $t\bar{t}$ l+jets background. Samples with $m_t = 170\,\text{GeV}$ and $m_t = 180\,\text{GeV}$ were used to estimate a 2σ variation in Run IIa. For Run IIb samples with $m_t = 170\,\text{GeV}$ and $m_t = 175\,\text{GeV}$ were used. This is then scaled to obtain a one standard deviation uncertainty, which is 2.5 GeV for Run IIa and 1.25 GeV for Run IIb. The difference between the Runs is due to the different MC samples used for the reconstruction and not due to any physical reason.

Luminosity An uncertainty of the measured luminosity of $\pm 6.1\,\%$ was taken into account [183].

Taggability Rate Functions The uncertainty on the Taggability Rate Function is calculated by raising and lowering the efficiencies associated with the Taggability by one standard deviation [124].

Tag Rate Functions The uncertainty on the light- , b-, and c-TRFs are calculated by raising and lowering the efficiencies associated with their appropriate TRF by one standard deviation [139, 140].

Heavy Flavor Scale Factor The uncertainty on the W+jets heavy flavor scale factor is 20%, according to [184], estimated by calculating the scale factor for different wokring points for the NN tagger.

b-Fragmentation The systematic uncertainties on the reweighting of the b-fragmentation function from between the default in PYTHIA and the tuned value was assumed to be the difference from the AOD tune to the SLD tune [164]. Therefore only a one sided systematic uncertainty is assigned.

Multijet Background The uncertainty of ϵ_{qcd} and ϵ_{sig} on the multijet yield calculated with the Matrix Method is taken into account as given in table 6.1.

Electron ID Scale Factor Systematic Uncertainties A systematic uncertainty of 2.5 % is used. This uncertainty includes the dependence of the electron ID scale factor on the following variables ignored in the parameterization. The contributions to the electron ID scale factor systematic uncertainties is estimated to come from the jet multiplicity dependence (1.2 %) of the H-matrix, track match and likelihood scale factor, and from dependences not taken into account (on p_T and ϕ of the em object) (2.2 %). The uncertainty due to the jet multiplicity dependence is estimated by calculating the difference between the efficiency when taking into account the dependence on ΔR between an electron and a jet. Without this dependence the scale factor uncertainty is 1.2 %.

Muon ID and Track Scale Factor Systematic Uncertainties According to the studies summarized in the muon ID certification note [102], uncertainties on muon ID and tracking are both 0.7 %.

Muon Isolation Scale Factor Systematic Uncertainties An uncertainty on the efficiency of the muon isolation criteria is extracted from [102] to be 2 %.

z Vertex Distribution Difference between Data and MC An uncertainty of 2.2 % on the difference between the z vertex distribution in simulated events and the data is used, as given in [185].

Primary Vertex Scale Factor The preselection efficiency is scaled with 98.9 % to correct for the difference of the primary vertex selection efficiency between data and Monte Carlo [169]. A relative uncertainty on this scale factor is estimated to be 1.5 % .

6.6.2 Systematic Uncertainty Comparison

A comparison of the nominal $M_{t\bar{t}}$ distributions to $+1\sigma$ and -1σ jet energy scale systematic variations for Run IIa and Run IIb are shown in figure 6.5 and figure 6.6 for the dominant $t\bar{t}$ background.

In tables 6.16 to 6.19, the relative change of the total number of predicted events in the Standard Model background and for a Z' boson with $m = 650\,\text{GeV}$ are listed for all systematic uncertainties. The numbers listed are for e+jets and μ+jets channels combined and reflect the influence of the systematic uncertainties on the estimated number of background events. For the multijet fake and lepton ID efficiencies, a weighted average of the electron and muon uncertainties has been quoted. The impact on the normalization and, in case of JES and others, also the differences in preselection efficiencies are taken into account. Tables 6.20 and 6.21 list the systematic uncertainties for one and $\geq$two b-tags. A value of < 0.01 means, that the absolute value of the uncertainty is below < 0.01, while 0.00 means the uncertainty is exactly zero. This is the case for the cross section, multijet fake rate, and the top quark mass variation for the resonances, as the resonance cross section is independent of those. The samples were produced with one top mass, as were all non standard model $t\bar{t}$ backgrounds, too. The slightly changed kinematics for the changed top mass has found to be minute in comparing Run IIa samples for 400 GeV and 700 GeV and can be neglected. Therefore not all samples have been reproduced for all masses.

It should be noted, that the tables do not take into account the shape change of the systematic uncertainties and can therefore not fully be interpreted as the systematic uncertainty in the limit for the resonance. The major difference between Run IIa and Run IIb are the increased JES and JER uncertainties for Run IIb which are not yet as precisely determined as for Run IIa. The jet identification has been increased due to the newly added silicon layer. While the amount of heavy flavor jets and luminosity uncertainties have been roughly equal, improvements in the multijet fake rate were made due to improved object definitions and analysis techniques. The top quark mass uncertainty has been improved, too.

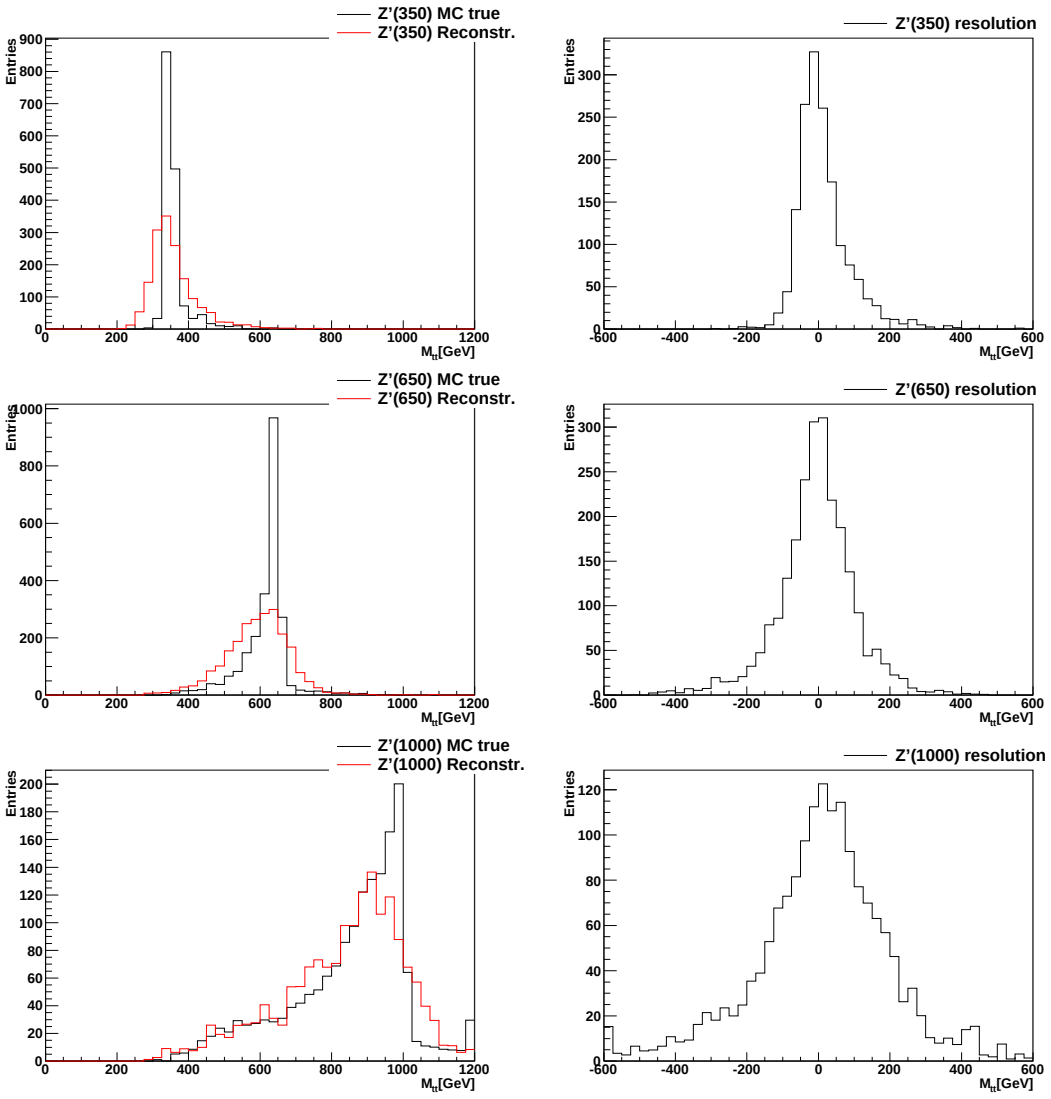


Figure 6.4: In the left plots, the Monte Carlo true as well as the reconstruction method information is used to calculate the invariant $t\bar{t}$ mass. The right side shows the mass difference between the two. The RMS of this distribution can then be used to compute the resolution.

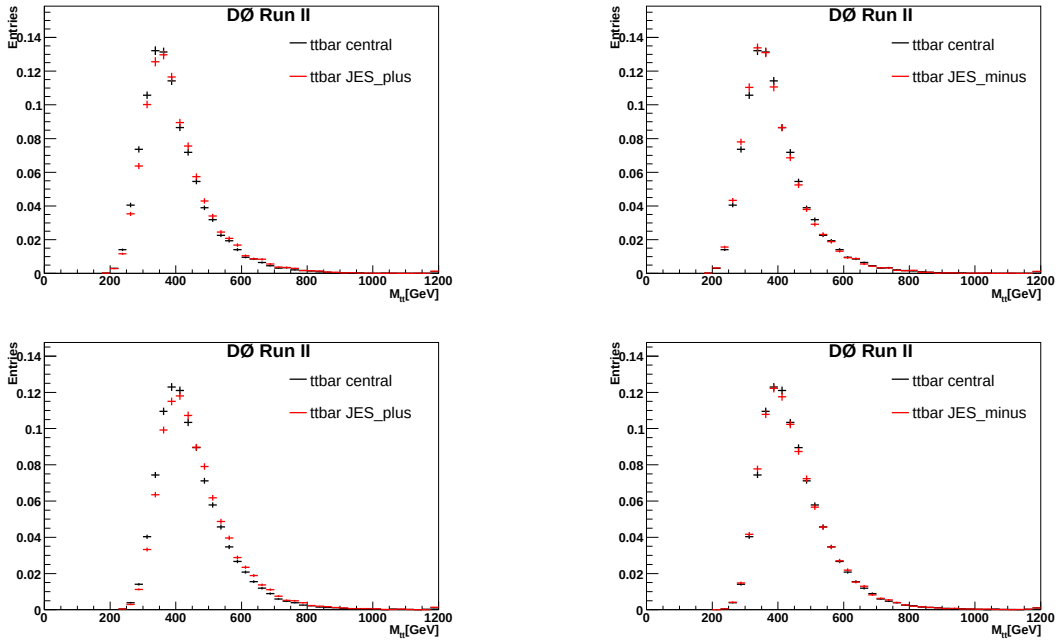


Figure 6.5: Run IIa shape comparison of the $M_{t\bar{t}}$ distributions of the nominal, $+1\sigma$ (right), and -1σ jet energy scale systematic uncertainty (left). The upper row corresponds to 3-jet events, the lower row corresponds to four or more jet events.

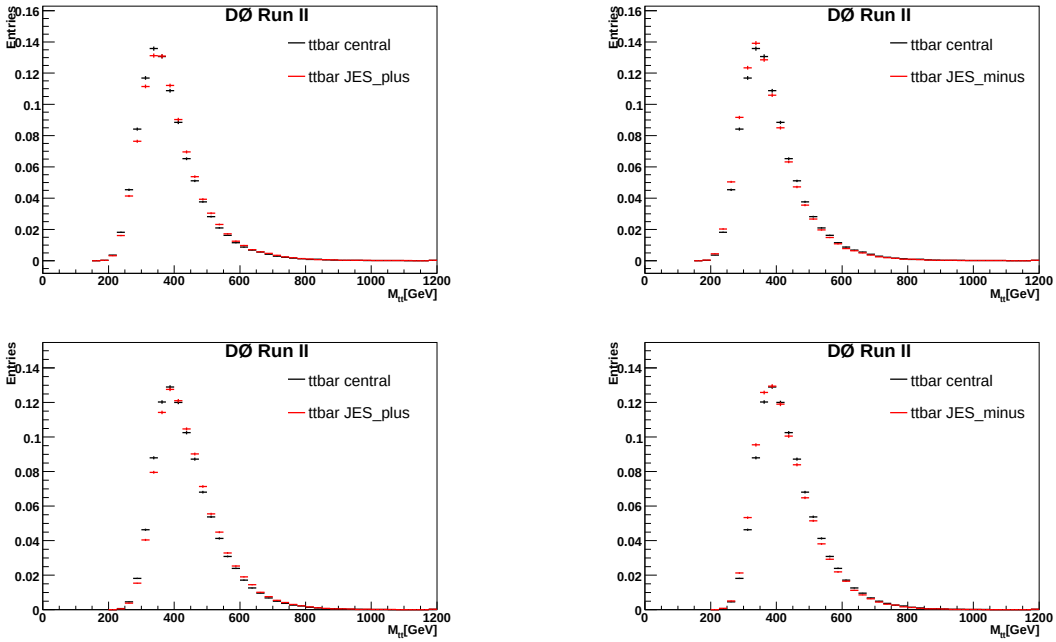


Figure 6.6: Run IIb shape comparison of the $M_{t\bar{t}}$ distributions of the nominal, $+1\sigma$ (right), and -1σ jet energy scale systematic uncertainty (left). The upper row corresponds to 3-jet events, the lower row corresponds to four or more jet events.

Source	rel. sys. uncertainty (%)							
	Standard Model processes (%)				Resonance $M_X = 650\,\text{GeV}$			
	3 jets		4 and more jets		3 jets		4 and more jets	
	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-
Jet Energy Scale	−0.34	0.36	−0.11	0.02	−1.99	1.46	3.23	−2.71
Jet Energy Resolution	−0.04	0.02	−0.02	< 0.01	< 0.01	0.27	< 0.01	0.09
Jet Identification	−0.12	0.12	−0.07	0.07	1.46	−1.46	−2.09	2.09
$\sigma_{t\bar{t}}(m_t = 172.4\,\text{GeV})$	1.93	−1.93	3.49	−3.48	0.00	0.00	0.00	0.00
b-tag TRF	1.07	−1.25	0.32	−0.65	1.06	−1.33	−0.09	−0.38
c-tag TRF	0.76	−0.77	0.26	−0.25	−0.07	0.07	−0.14	0.14
light-tag TRF	0.81	−0.81	0.34	−0.34	< 0.01	< 0.01	−0.05	0.05
Taggability	1.05	−1.08	0.68	−0.73	0.86	−0.93	0.43	−0.52
Heavy Flavor Scale Factor for W+jets	1.80	−1.69	0.79	−0.75	0.00	0.00	0.00	0.00
b-fragmentation	0.16	< 0.01	0.11	< 0.01	0.93	< 0.01	1.42	< 0.01
Multijet Lepton Fake Rate	0.75	−0.66	−0.19	0.16	0.00	0.00	0.00	0.00
Luminosity	2.27	−2.02	3.65	−3.35	6.10	−5.77	6.08	−5.79
Top Mass Variation	1.16	1.93	0.33	1.12	0.00	0.00	0.00	0.00
Selection Efficiencies (incl. Lepton ID)	1.57	−1.57	2.65	−2.64	3.58	−3.65	3.61	−3.61

Table 6.16: Summary of the relative systematic change on the overall normalization of the Standard Model background and for a resonance mass of $M_X = 650\,\text{GeV}$ with 1 b-tag for Run IIa.

Source	rel. sys. uncertainty (%)							
	Standard Model processes (%)				Resonance $M_X = 650\,\text{GeV}$			
	3 jets		4 and more jets		3 jets		4 and more jets	
	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-
Jet Energy Scale	−0.71	−0.65	−0.40	−1.20	−1.93	1.12	3.15	−2.71
Jet Energy Resolution	−0.80	−0.72	−0.79	−0.79	−0.16	0.32	< 0.01	0.09
Jet Identification	−1.47	1.47	−1.15	1.15	0.64	−0.64	−2.27	2.27
$\sigma_{t\bar{t}}(m_t = 172.4\,\text{GeV})$	3.79	−3.79	5.22	−5.22	0.00	0.00	0.00	0.00
b-tag TRF	9.06	−8.63	8.96	−8.57	9.63	−9.15	9.27	−8.83
c-tag TRF	0.75	−0.72	0.72	−0.73	0.80	−0.64	0.61	−0.70
light-tag TRF	0.31	−0.30	0.26	−0.26	0.16	−0.16	0.17	−0.26
Taggability	3.02	−2.93	2.98	−2.92	3.37	−3.37	3.23	−3.23
Heavy Flavor Scale Factor for W+jets	1.28	−1.21	0.33	−0.33	0.00	0.00	0.00	0.00
b-fragmentation	0.30	< 0.01	0.14	< 0.01	0.64	< 0.01	1.05	< 0.01
Multijet Lepton Fake Rate	0.16	−0.13	−0.04	0.01	0.00	0.00	0.00	0.00
Luminosity	4.30	−3.99	5.33	−4.99	6.10	−5.78	6.12	−5.77
Top Mass Variation	3.01	3.59	0.95	1.33	0.00	0.00	0.00	0.00
Selection Efficiencies (incl. Lepton ID)	2.78	−2.76	3.33	−3.34	3.69	−3.53	3.58	−3.58

Table 6.17: Summary of the relative systematic change on the overall normalization of the Standard Model background and for a resonance mass of $M_X = 650\,\text{GeV}$ with 2 b-tags for Run IIa.

Source	rel. sys. uncertainty (%)							
	Standard Model processes (%)				Resonance $M_X = 650\,\text{GeV}$			
	3 jets		4 and more jets		3 jets		4 and more jets	
	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-
Jet Energy Scale	−1.09	1.19	2.41	−2.20	−1.45	0.95	3.59	−3.65
Jet Energy Resolution	−3.05	1.56	0.44	−0.56	−1.89	0.70	1.96	−1.33
Jet Identification	0.16	−0.16	−1.00	1.00	0.27	−0.27	−1.50	1.50
$\sigma_{t\bar{t}}(m_t = 172.4\,\text{GeV})$	2.17	−2.79	4.08	−5.24	0.00	0.00	0.00	0.00
b-tag TRF	−1.43	−4.39	−0.82	−2.37	0.45	−3.48	0.00	−1.90
c-tag TRF	−3.38	−5.42	−2.07	−2.68	−3.50	−3.48	−2.00	−1.81
light-tag TRF	−0.28	−2.80	−0.50	−1.39	−1.36	−1.34	−0.87	−0.72
Taggability	1.45	−1.54	0.82	−0.94	1.18	−1.34	0.67	−0.80
Heavy Flavor Scale Factor for W+jets	1.04	0.04	0.45	−0.17	0.00	0.00	0.00	0.00
b-fragmentation	−5.39	< 0.01	−2.67	< 0.01	−3.48	< 0.01	−1.81	< 0.01
Multijet Lepton Fake Rate	−0.20	0.19	−0.16	0.15	0.00	0.00	0.00	0.00
Luminosity	1.78	−1.68	3.33	−3.14	6.11	−5.75	6.10	−5.74
Top Mass Variation	−0.85	1.01	−1.18	1.78	0.00	0.00	0.00	0.00
Selection Efficiencies (incl. Lepton ID)	1.14	−1.14	2.01	−2.01	3.64	−3.61	3.62	−3.60

Table 6.18: Summary of the relative systematic change on the overall normalization of the Standard Model background and for a resonance mass of $M_X = 650\,\text{GeV}$ with 1 b-tag for Run IIb.

Source	rel. sys. uncertainty (%)							
	Standard Model processes (%)				Resonance $M_X = 650\,\text{GeV}$			
	3 jets		4 and more jets		3 jets		4 and more jets	
	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-
Jet Energy Scale	−1.61	1.72	3.78	−3.81	−2.50	1.44	3.55	−3.48
Jet Energy Resolution	−3.72	2.13	0.79	−0.76	−2.94	0.81	1.11	−0.73
Jet identification	−0.17	0.17	−1.79	1.79	−0.06	0.06	−1.91	1.91
$\sigma_{t\bar{t}}(m_t = 172.4\,\text{GeV})$	4.68	−6.03	6.38	−8.22	−0.00	0.00	−0.00	0.00
b-tag TRF	5.40	−13.77	5.58	−13.36	5.07	−14.76	4.94	−13.94
c-tag TRF	−12.82	−14.69	−12.52	−14.19	−14.13	−15.38	−13.21	−14.71
light-tag TRF	−3.15	−4.44	−3.17	−4.12	−4.32	−5.13	−3.79	−4.73
Taggability	3.84	−3.79	3.69	−3.65	4.38	−4.75	4.14	−4.24
Heavy Flavor Scale Factor for W+jets	0.86	0.04	0.23	−0.09	−0.00	0.00	−0.00	0.00
b-fragmentation	−14.63	< 0.01	−14.18	< 0.01	−15.38	< 0.01	−14.71	< 0.01
Multijet Lepton Fake Rate	−0.23	0.21	−0.07	0.06	−0.00	0.00	−0.00	0.00
Luminosity	3.81	−3.60	5.20	−4.91	6.13	−5.82	6.08	−5.77
Top Mass Variation	−1.65	1.94	−1.82	2.54	0.00	0.00	0.00	0.00
Selection Efficiencies (incl. Lepton ID)	2.39	−2.40	3.13	−3.13	3.63	−3.63	3.62	−3.62

Table 6.19: Summary of the relative systematic change on the overall normalization of the Standard Model background and for a resonance mass of $M_X = 650\,\text{GeV}$ with 2 b-tags for Run IIb.

Source	rel. sys. uncertainty (%)							
	Standard Model processes (%)				Resonance $M_X = 650\,\text{GeV}$			
	3 jets		4 and more jets		3 jets		4 and more jets	
	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-
Jet Energy Calibration	−0.41	0.18	−0.19	−0.34	−2.02	1.31	3.20	−2.74
Jet Energy Resolution	−0.18	−0.11	−0.24	−0.24	−0.09	0.23	−0.03	0.09
Jet Identification	−0.36	0.36	−0.39	0.39	1.17	−1.17	−2.15	2.15
$\sigma_{t\bar{t}}(m_t = 170\,\text{GeV})$	2.27	−2.27	4.00	−3.99	0.00	0.00	0.00	0.00
W+jets (heavy flavor)	1.70	−1.61	0.65	−0.62	0.00	0.00	0.00	0.00
b-fragmentation	0.18	< 0.01	0.12	< 0.01	0.80	< 0.01	1.29	< 0.01
Multijet Lepton Fake Rate	0.64	−0.57	−0.15	0.12	0.00	0.00	0.00	0.00
Luminosity	2.63	−2.37	4.14	−3.83	6.05	−5.77	6.09	−5.75
Top Mass Variation	1.49	2.23	0.51	1.19	0.00	0.00	0.00	0.00
Selection Efficiencies (incl. Lepton ID)	1.79	−1.79	2.85	−2.84	3.56	−3.66	3.60	−3.63
b-tagging	3.05	−3.13	3.21	−3.32	3.85	−4.06	3.50	−3.71

Table 6.20: Summary of the relative systematic change on the overall normalization of the Standard Model background and for a resonance mass of $M_X = 650\,\text{GeV}$ with ≥ 1 b-tag for Run IIa.

Source	rel. sys. uncertainty (%)							
	Standard Model processes (%)				Resonance $M_X = 650\,\text{GeV}$			
	3 jets		4 and more jets		3 jets		4 and more jets	
	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-	σ^+	σ^-
Jet Energy Scale	−1.18	1.28	2.81	−2.67	−1.72	1.08	3.58	−3.60
Jet Energy Resolution	−3.17	1.66	0.54	−0.62	−2.18	0.75	1.68	−1.13
Jet Identification	0.11	−0.11	−1.23	1.23	0.18	−0.18	−1.63	1.63
$\sigma_{t\bar{t}}(m_t = 170\,\text{GeV})$	2.60	−3.34	4.76	−6.11	−0.00	0.00	0.00	0.00
W+jets (heavy flavor)	1.01	0.04	0.39	−0.15	−0.00	0.00	0.00	0.00
b-fragmentation	−6.97	< 0.01	−6.05	< 0.01	−6.65	< 0.01	−6.07	< 0.01
Multijet Lepton Fake Rate	−0.22	0.20	−0.14	0.13	−0.00	0.00	0.00	0.00
Luminosity	2.13	−2.01	3.88	−3.66	6.10	−5.75	6.10	−5.75
Top Mass Variation	−0.99	1.18	−1.36	2.00	0.00	0.00	0.00	0.00
Selection Efficiencies (incl. Lepton ID)	1.36	−1.35	2.34	−2.34	3.62	−3.62	3.61	−3.61
b-tagging	5.39	−9.92	5.65	−8.71	7.19	−9.83	6.45	−8.91

Table 6.21: Summary of the relative systematic change on the overall normalization of the Standard Model background and for a resonance mass of $M_X = 650\,\text{GeV}$ with ≥ 1 b-tags for Run IIb.

6.7 Limit Calculation

Physical measurements have uncertainties. Results consistent with zero, are in general presented as an upper limit on the physical quantity under consideration. Usually then the confidence level is chosen to be 95 %. Often physical quantities or observables have boundaries, like the cross section. This value is constrained to be positive by the physical process though the measured value might turn out to be negative. There are several possible ways how to treat those cases.

Over time, two major schools have developed on how to cope with statistical processes: the Frequentist and the Bayesian school. Both use different ways to calculate a limit. They do not give the same numerical result even when employing the same set of input data. However, they also do not give the same answer as they use different ways to describe the result. Both methods have their drawbacks and advantages. For this analysis the Bayesian approach was chosen, which is described briefly in this section, followed by the application of this method in the analysis including some alterations needed to cope with certain aspects of the analysis.

6.7.1 The Bayesian Approach

The $t\bar{t}$ invariant mass spectrum, as shown in figure 7.1, is used to perform a binned likelihood fit of the signal and background expectations compared to data. The backgrounds, including Standard Model $t\bar{t}$ production, are normalized to the predictions, as described in section 6.2. A Bayesian approach leads to upper limits on $\sigma_X \times B(X \to t\bar{t})$ for the different resonance masses.

A Poisson distribution is assumed for the number of observed events in each bin, as well as flat prior probabilities for the signal cross sections between 0 and 10 pb. The chosen upper limit for the prior is sufficiently large to have no impact on the limit calculation. More information about the limit calculation can be found in the single top analyses [186, 187].

The likelihood to observe an event count D, given the expectation value is d, is given by the Poisson distribution:

$$L(D|d) = \frac{e^{-d}d^D}{D!} \ . \tag{6.8}$$

In the case we do have Monte Carlo simulated events, this equation needs to be expanded to non-integer values giving:

$$L(D|d) = \frac{e^{-d}d^D}{\Gamma(D+1)} \ , \tag{6.9}$$

where Γ is the Gamma function.

The mean count d is the sum of the predicted contributions from the signal and N background sources:

$$d = \alpha\mathcal{L}\sigma + \sum_{i=1}^{N} b_i = a\sigma + \sum_{i=1}^{N} b_i \ , \tag{6.10}$$

where α is the signal acceptance, $\mathcal{L}$ the luminosity and σ the cross section of the process of interest, b_i the yield for the background process i and $a = \alpha\mathcal{L}$.

In case more than a counting experiment is done, a binned likelihood fit of the signal and background expectations can be done by expanding the formula to:

$$L(\vec{D}|\vec{d}) = L(\vec{D}|\sigma, \vec{a}, \vec{b}) = \prod_{j=1}^{M} L(D_j|d_j) \ , \tag{6.11}$$

where $\vec{D}$ and $\vec{d}$ are the vectors of the observed and expected counts in M bins, and $\vec{a}$ and $\vec{b}$ are the vectors of the effective luminosity and background contribution. This can be further expanded to any number of signal channels, as long as the channels are independent from each other. The combined likelihood is then a product of the single likelihoods for each channel. In this analysis one signal is assumed, while taking 16 different channels into account. The analysis channels are a combination of Run IIa / Run IIB, e+jets / μ+jets, one b-tag / two $b-$tags, and three and four jet bin. This is done for each of the resonance samples separately.

From Bayes' theorem the posterior density $p(\sigma, \vec{a}, \vec{b}|\vec{D})$ can be computed [188]:

$$p(\sigma|\vec{D}) = \frac{1}{N} \int\int L(\vec{D}|\sigma, \vec{a}, \vec{b}) \pi(\sigma, \vec{a}, \vec{b}) d\vec{a} d\vec{b} \ . \tag{6.12}$$

The parameter N is an overall normalization obtained using $\int_0^{\sigma_{\max}} p(\sigma|D)d\sigma = 1$, where $\sigma_{\max}$ is a value, where the integration numerically ends and the posterior is practically zero. Furthermore the independence of the signal cross section σ and the parameters $\vec{a}$ and $\vec{b}$ is assumed, factorizing the prior:

$$\pi(\sigma, \vec{a}, \vec{b}) = \pi(\vec{a}, \vec{b})\pi(\sigma) \ . \tag{6.13}$$

The prior density for the cross section $\pi(\sigma)$ is assumed to be flat in σ:

$$\begin{aligned} \pi(\sigma) &= \frac{1}{\sigma_{\max}}, \ 0 < \sigma < \sigma_{\max} \\ &= 0, \ \text{otherwise} \end{aligned} \tag{6.14}$$

The choice of a flat prior for the cross section is rather a convention in high energy physics and not a mathematically or physically proven fact. In this analysis the prior has $\sigma_{\max} = 10\,\text{pb}$, which is sufficiently large, as all posterior densities turned out to be minute at $\sigma_{\max}$. The posterior probability density can be calculated by:

$$p(\sigma|\vec{D}) = \frac{1}{N\sigma_{\max}} \int\int L(\vec{D}|\sigma,\vec{a},\vec{b})\pi(\vec{a},\vec{b})d\vec{a}d\vec{b} \; . \tag{6.15}$$

with a Bayesian upper limit σ_{CL} at a confidence level CL:

$$\int_0^{\sigma_{\text{CL}}} p(\sigma|\vec{D}) = \text{CL} \; . \tag{6.16}$$

The cross section of the process can be extracted from the posterior probability density in different ways, like the mean, median or mode. The mode or peak of the distribution is chosen as the measured cross section and the 68 % probability interval as the estimate of its uncertainty. Ensemble tests done in [189] showed that this, though Bayesian, has approximatly 68 % coverage probability and can therefore interpreted as frequentist intervals, if desired.

The integration in equation 6.15 is done numerically using Monte Carlo importance sampling. A large number K of points $(\vec{a}_K, \vec{b}_K)$ are randomly sampled from $\pi(\vec{a},\vec{b})$ and the posterior density is estimated by:

$$\begin{aligned} p(\sigma|\vec{D}) &\propto \int\int L(\vec{D}|\sigma,\vec{a},\vec{b})\pi(\vec{a},\vec{b})d\vec{a}d\vec{b} \; , \\ &\approx \frac{1}{K}\sum_{k=1}^{K} L(\vec{D}|\sigma,\vec{a}_k,\vec{b}_k) \end{aligned} \tag{6.17}$$

6.7.2 Systematic Uncertainty Treatment

The prior $\pi(\vec{a},\vec{b})$ encodes the knowledge of the effective signal luminosities and the background yields. The method used in this analysis is called direct sampling, which evaluates the uncertainty by drawing random numbers from a Gaussian distribution with width one and mean zero $g(0,1)_{\text{isys}}$ for each systematic uncertainty i. For each systematic uncertainty a shift Δy_{isys} is assigned. For this $2n+1$ number of templates are needed where n is the number of systematic uncertainties: the central template plus an upward and downward template for each systematic uncertainty. If the Gaussian random number is positive, the upward systematic uncertainty variation is used and the downward otherwise. The shift

for each bin is calculated by:

$$\Delta y_{\text{isys}} = s^{+}_{\text{tot}} \times g(0,1)_{\text{isys}} \times (s^{+}_{\text{isys}} - y), \text{ if} g(0,1)_{\text{isys}} > 0 \quad (6.18)$$

$$= s^{-}_{\text{tot}} \times g(0,1)_{\text{isys}} \times (y - s^{-}_{\text{isys}}), \text{ otherwise.} \quad (6.19)$$

The value s^{+}_{isys} (s^{-}_{isys}) is the upward (downward) systematic uncertainty template entry value, y the value of the central template, and $s^{\pm}_{\text{tot}}$ a total normalization factor normally set to one. This is done for each bin of the corresponding signal (background) source and channel. This changes the value in a given bin from a source from y to y':

$$y' = y + \sum_{\text{isys}} \Delta y_{\text{isys}} \quad (6.20)$$

A set of this Gaussian random numbers needs to be drawn for each systematic uncertainty. The integration over the Gaussian is performed by drawing random numbers from the Gaussian many times and computing a likelihood for each sample. The final posterior is then the sum of the individual systematic uncertainty likelihoods.

For numerical reasons, two specific cases need to be dealt with. First, the resulting entry in the bin may never be negative. If the sum of a shift returns a negative value, the shift is set to $-y$ resulting in a zero bin entry. If this happens more often than a few percent the Gaussian set is discarded and a new one is drawn. Secondly, the same is done, when the sum of all shifts return a negative value. Also some small corrections are made due to limited Monte Carlo statistics and large shifts with more than 100 % the value of y. Details of this can be found in [189].

In order to save computing time the posterior cross section density is evaluated at five points, which are N times the expected cross section and $N \in \{0, 1, 2, 3, 4\}$. The posterior density is then estimated from 1000 samples at these points. The samples are sorted and additional new samples are only kept, if the posterior density exceeds a certain threshold of at least one of these points. This saves computing time during the integration while not biasing the result.

6.7.3 Ensemble Testing

When applied to real data $\vec{D}$ the posterior densities should look reasonable. To cross check the behavior a set of pseudo experiments is done, ensembles are generated using the same procedure described above. For each pseudo data set a systematic uncertainty sample is chosen at random and a count from each bin is sampled with a Poisson distribution, with a mean of the background yield in that sample in this bin. This is treated as the data input.

The limits are calculated in the same way as described above taking all correlations into account. The resulting posterior densities are then stored and a mean value and uncertainty is calculated. These are then used as a cross check. In case of this analysis the ensembles including a $1\,\sigma$ uncertainty should form a band around the expected limit, as the signal in the pseudo experiments is assumed to be zero. This can be seen in section 7.3.

6.7.4 Testing of the `top_statistics` Framework with a 2D-Fit

The program used has been extensively tested, also with the possibility to make two-dimensional fits. The $t\bar{t}$ pair production cross section and a Z' boson sample were used. Various cross sections have been assumed for the $t\bar{t}$ pair production cross section, while keeping the Z' boson cross section in the pseudo data zero. The program reconstructed both cross sections correctly well within the $1\,\sigma$ range.

6.7.5 Signal Acceptance Correction

No $t\bar{t}$ resonance contributions were assumed when estimating the W+jets background. But when calculating the limits, the presence of the resonance has to be taken into account, since in this case the number of estimated W+jets events would be reduced. This is due to the fact, that in addition to the backgrounds normalized to their cross-section (SM top, Diboson, Z+jets, ...), the expected resonance events also have to be subtracted from $N^{\text{untag}}_{W+\text{top}}$ obtained from the Matrix Method in equation 6.2. To correctly treat the W+jets background in the presence of a resonance, the following considerations were made:

- The untagged data sample consists of $N^{\text{untag}}_{W+\text{top}+X}$ and $N^{\text{untag}}_{\text{QCD}}$:

$$N^{\text{untag}}_{\text{data}} = N^{\text{untag}}_{W+\text{top}+X} + N^{\text{untag}}_{\text{QCD}}. \tag{6.21}$$

- $N^{\text{untag}}_{W+\text{top}+X}$ is the sum of the W+jets, other SM backgrounds with real leptons (dominated by SM top, but including Diboson and Z+jets) and resonance contributions. So the W+jets contribution is calculated by:

$$N^{\text{untag}}_{W+jets} = N^{\text{untag}}_{W+\text{top}+X} - N^{\text{untag}}_{\text{top}} - N^{\text{untag}}_{X}. \tag{6.22}$$

- This leads to:

$$N^{\text{untag}}_{\text{data}} = (N^{\text{untag}}_{W+\text{top}+X} - N^{\text{untag}}_{\text{top}} - N^{\text{untag}}_{X}) + N^{\text{untag}}_{\text{top}} + N^{\text{untag}}_{X} + N^{\text{untag}}_{\text{QCD}}. \tag{6.23}$$

- After applying b-tagging, the number of untagged events for each sample is multiplied with the b-tag efficiency in that particular sample, which is then the number of events in the tagged sample.

$$N_{\text{data}}^{\text{tag}} = \varepsilon_b^W (N_{W+\text{top}+X}^{\text{untag}} - N_{\text{top}}^{\text{untag}} - N_X^{\text{untag}}) + \varepsilon_b^{top} N_{\text{top}}^{\text{untag}} + \varepsilon_b^X N_X^{\text{untag}} + \varepsilon_b^{\text{QCD}} N_{\text{QCD}}^{\text{untag}}. \tag{6.24}$$

- Solving this equation for the number of resonance events N_X gives:

$$(\varepsilon_b^X - \varepsilon_b^W) N_X^{\text{untag}} = N_{\text{data}}^{\text{tag}} - \varepsilon_b^{\text{top}} N_{\text{top}}^{\text{untag}} - \varepsilon_b^W (N_{W+\text{top}+X}^{\text{untag}} - N_{\text{top}}^{untag}) - \varepsilon_b^{\text{QCD}} N_{\text{QCD}}^{\text{untag}}. \tag{6.25}$$

Thus, instead of correcting the number of W+jets events for the limit calculation, the resonance input can be corrected for the b-tag efficiency in a W+jets sample. This was done for all resonance samples and also for all systematic uncertainties which potentially change the b-tag efficiency. This method was validated in [156].

CHAPTER 7

Cross Section Limits for Resonances and Mass Limits for Z' Bosons

After the invariant $M_{t\bar{t}}$ distribution is calculated, the cross section can be calculated. Is the measured cross section either zero or within about three standard deviation to zero, no signal can be claimed to be seen. The following chapter will summarize the results, which are computed with the analysis method described in the last chapter including the systematic uncertainties and the extension to any small width resonance besides the chosen Z' benchmark model used for the mass comparison in the Tevatron experiments. A method to extract the top quark mass using the invariant mass distribution is briefly described in appendix A, as no studies of systematic uncertainties were made for this analysis.

The next sections describe the different forms of limits, which can be obtained. The limits are calculated with the procedure described in section 6.7. First, the expected limit will be described, which tests the SM hypothesis, and can be interpreted as the sensitivity of the experiment. The observed limit is given by the measured data and describes the real measured limit. Values below the expected one show a less than expected number of events than in the SM and favor the exclusion, while a very high value might give hint for a signal. The measured cross section will be shown, too. It can be seen, that no signal can be claimed. The method is tested with ensemble tests made of pseudo experiments, which give the range in which the expected limit should lie. Finally, the result on the vector and axial vector resonances is discussed, widening the benchmark test of the Z' boson to a more general small width resonance.

7.1 Expected Limits

The expected limit is calculated by assuming a SM like data distribution and applying the limit setting method to it. The previously discussed backgrounds are fitted with the resonance signal depending on its mass to the data. This model uses the SM only hypothesis and uses the expected distribution as pseudo data. This gives an estimate for the expected sensitivity of the method.

The invariant mass distribution for the 3-jet and 4-jet bins are shown in figure 7.1. Corresponding plots for the individual channels can be found in figures C.54 and C.55 in appendix C.

	3 jets	≥ 4 jets
$t\bar{t}$	624	721
Single Top	47	13
Diboson	32	8
W+jets	592	129
Z+jets	85	26
Multijet	84	22
Total Background	1464	919
Data	1411	934

Table 7.1: Number of observed and predicted events.

Expected limits were computed by using the SM expectation as input data in the limit calculation with statistical uncertainties only, as well as including full systematic uncertainties. They were computed for various analysis options discussed below. The option to be applied to real data was chosen based on the best expected performance.

In computing the limit, the distributions for one and two or more b-tags can be handled as separate channels or as a combined one of one or more b-tag channels. The two methods give about the same results as was shown in [156]. For the final computation the separate channels were used, as they give a slightly better performance, as more input information can be used. Following the results of the comparisons made in this section, the final result will be computed using the medium b-tag working point and the direct $M_{t\bar{t}}$ reconstruction.

To display the varying contribution of some of the systematics over the range of resonance masses, figure 7.2 compares the expected limit without systematics, with partial systematics, and with full systematics. It can be seen, that the top quark mass uncertainty dominates the uncertainty at low resonance masses,

where the shapes of the SM production and the resonant production have similar p_T spectra. At larger resonance masses, the limit becomes statistically limited.

7.2 Observed Results

After all selection cuts, 382 events in the e+jets channel and 326 events in the μ+jets channel for Run IIa. 911 events remain in the e+jets channel and 726 events in the μ+jets channel for Run IIb. Invariant mass distributions are computed for events with exactly one b-tag and for those with more than one b-tag. Additionally, the distributions are separated into three jet and four or more jet events. Corresponding plots for the individual channels can be found in Figures C.54 and C.55 in Appendix C.

The expected and observed 95% C.L. upper limits on $\sigma_X \times B(X \to t\bar{t})$, as a function of M_X, are summarized in tables 7.2 to 7.4 and displayed in figure 7.3 to figure 7.5. The expected limits with 68 % C.L. and 95 % C.L. are displayed as a shaded aera. The observed limit is a solid red line, while the measured cross section inlcuding its uncertainty is a dark red line. This figure also includes the predicted $\sigma_X \times B(X \to t\bar{t})$ for a leptophobic Z' boson with $\Gamma_{Z'} = 0.012 M_{Z'}$ computed, using the CTEQ6L1 parton distribution function with LO k-factors as a blue dashed line. Combining the information of the events with four or more jets and those with three jets, the expected limit on the Z' mass is 880 GeV. The observed limits exclude a $M_{Z'} < 820$ GeV at 95% C.L. in the combined channels.

	Statistical Uncertainty		Stat. and Syst. Uncertainty	
M_X [GeV]	Exp. Limit [pb]	Obs. Limit [pb]	Exp. Limit [pb]	Obs. Limit [pb]
350	1.10	2.87	1.83	4.56
400	1.07	2.36	1.89	2.97
450	0.94	1.84	1.48	1.92
500	0.78	1.36	1.08	1.29
550	0.64	1.08	0.79	0.93
600	0.51	1.04	0.57	0.92
650	0.44	1.04	0.46	1.00
750	0.31	0.72	0.31	0.64
850	0.24	0.52	0.24	0.44
1000	0.19	0.34	0.19	0.28

Table 7.2: Expected and observed limits on $\sigma_X \times B(X \to t\bar{t})$ at the 95% confidence level when combining all channels with and without systematic uncertainties taken into account for Run IIa.

	Statistical Uncertainty		Stat. and Syst. Uncertainty	
M_X [GeV]	Exp. Limit [pb]	Obs. Limit [pb]	Exp. Limit [pb]	Obs. Limit [pb]
350	0.66	0.26	1.26	0.78
400	0.67	0.27	1.49	0.76
450	0.59	0.36	1.32	1.06
500	0.51	0.42	1.02	1.16
550	0.42	0.41	0.70	0.90
600	0.34	0.39	0.50	0.74
650	0.28	0.35	0.38	0.55
750	0.20	0.19	0.25	0.22
850	0.16	0.20	0.18	0.23
1000	0.12	0.18	0.14	0.20

Table 7.3: Expected and observed limits on $\sigma_X \times B(X \to t\bar{t})$ at the 95% confidence level when combining all channels with and without systematic uncertainties taken into account for Run IIb.

For the Run IIa data the limits are shown in figure 7.3. A $2\,\sigma$ upward fluctuation for a 350 GeV Z'-boson for the 4-jet bin limit as well as the combined (3-jet and 4-jet) limit can be seen. The 3-jet bin limit shows an excess, but it is below the $2\,\sigma$ contour. As there is no evidence in Run IIb as seen in figure 7.4 for such an excess, the excess most likely is a statistical fluctuation and we cannot claim evidence for a Z' resonance at 350 GeV.

In Run IIa the invariant mass region of $600 - 750$ GeV for both the events with three and those with four or more jets, as well as in Run IIb events with four or more jets an excess above zero is measured, as can seen in figures 7.3 and 7.4. This also is visible on the combined result in figure 7.5. The limit for events with three jets even extends the above zero measurement upward. This leads to an about 30 % worse limit than expected. As the measured cross section is always consistent with zero within two standard deviations, one could not claim a hint of a resonance. The non zero contribution of the cross section in the higher mass region also changes the slope of the observed cross section limit, leading to a and therefore worse limit.

In figures 7.7–7.10 the posterior probability densities for a Z' boson with a mass of 650 GeV are shown as a function of $\sigma_X \times B(X \to t\bar{t})$ for the expected and observed data. The chosen upper limit for 10 pb for the prior is sufficiently high, that it cannot influence the result, as all posteriors are minute at $3 - 5$ pb. For illustration purposes, the $t\bar{t}$ invariant mass distribution is shown with an assumed resonance contribution, where the resonance cross section is set to the observed limit including systematics. The legend contains the total number of events for each Monte Carlo sample and the observed data.

	Statistical Uncertainty		Stat. and Syst. Uncertainty	
M_X [GeV]	Exp. Limit [pb]	Obs. Limit [pb]	Exp. Limit [pb]	Obs. Limit [pb]
350	0.56	0.41	0.98	1.00
400	0.56	0.38	1.19	0.87
450	0.50	0.48	1.01	0.96
500	0.42	0.49	0.72	0.81
550	0.35	0.45	0.49	0.61
600	0.28	0.46	0.35	0.56
650	0.23	0.45	0.27	0.53
750	0.16	0.26	0.18	0.24
850	0.12	0.23	0.13	0.22
1000	0.09	0.17	0.10	0.16

Table 7.4: Expected and observed limits on $\sigma_X \times B(X \to t\bar{t})$ at the 95% confidence level when combining all channels with and without systematic uncertainties taken into account for full Run II.

Using as a benchmark model the model of a leptophobic topcolor assisted technicolor Z' boson, as described in section 2.3.4, such a resonance can be excluded with $M_{Z'} < 820$ GeV and width $\Gamma_{Z'} = 0.012 M_{Z'}$ at 95% C.L. In figure 7.5, the observed limits at the 95% C.L. were already shown together with the predicted cross section of a topcolor Z' boson for a width of $\Gamma_{Z'} = 0.012 M_{Z'}$. The final result can also be seen in a logarithmic plot shown in figure 7.6.

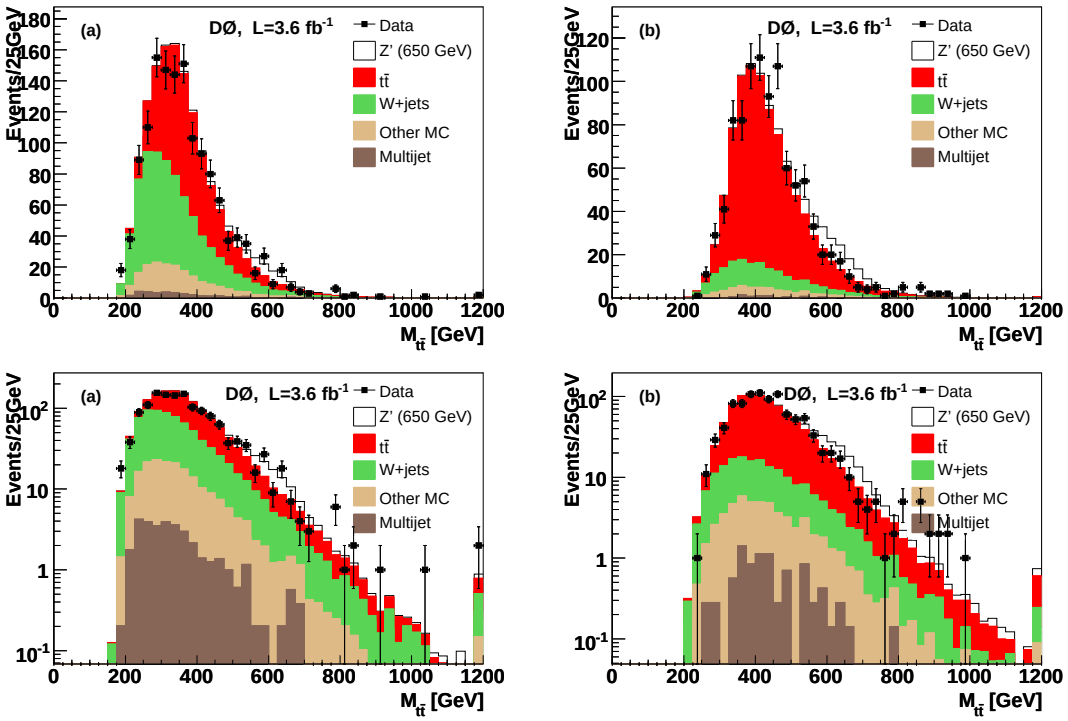


Figure 7.1: Expected and observed $t\bar{t}$ invariant mass distribution for the combined ℓ + 3jets, and ℓ + 4 or more jets channels, with at least one identified b-jet. The left column (a) shows the 3-jet bin, while the right one (b) shows the 4-jet bin with at least one b-tag. The first row has a linear scale and the second a logarithmic one. The error bars for the data drawn on top of the SM background indicate the statistical uncertainty. Superimposed as white area is the theory signal for a top-color-assisted technicolor Z' boson with $M_{Z'} = 650\,\text{GeV}$ for illustration purposes. The number of data, signal and expected background events from each source are indicated in table 7.1.

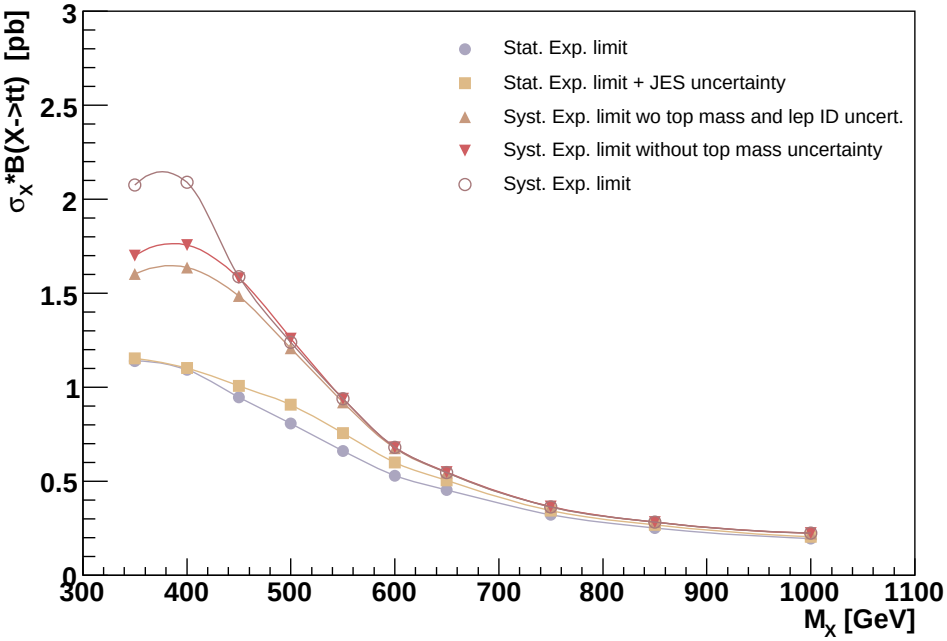


Figure 7.2: Expected limits without systematic uncertainties (circles), including only JES systematic uncertainty (squares), systematic uncertainties except selection efficiencies, m_t and luminosity (upward pointing triangles), all systematic uncetainties except m_t (downward pointing triangles) and complete systematic uncertainties (open circles).

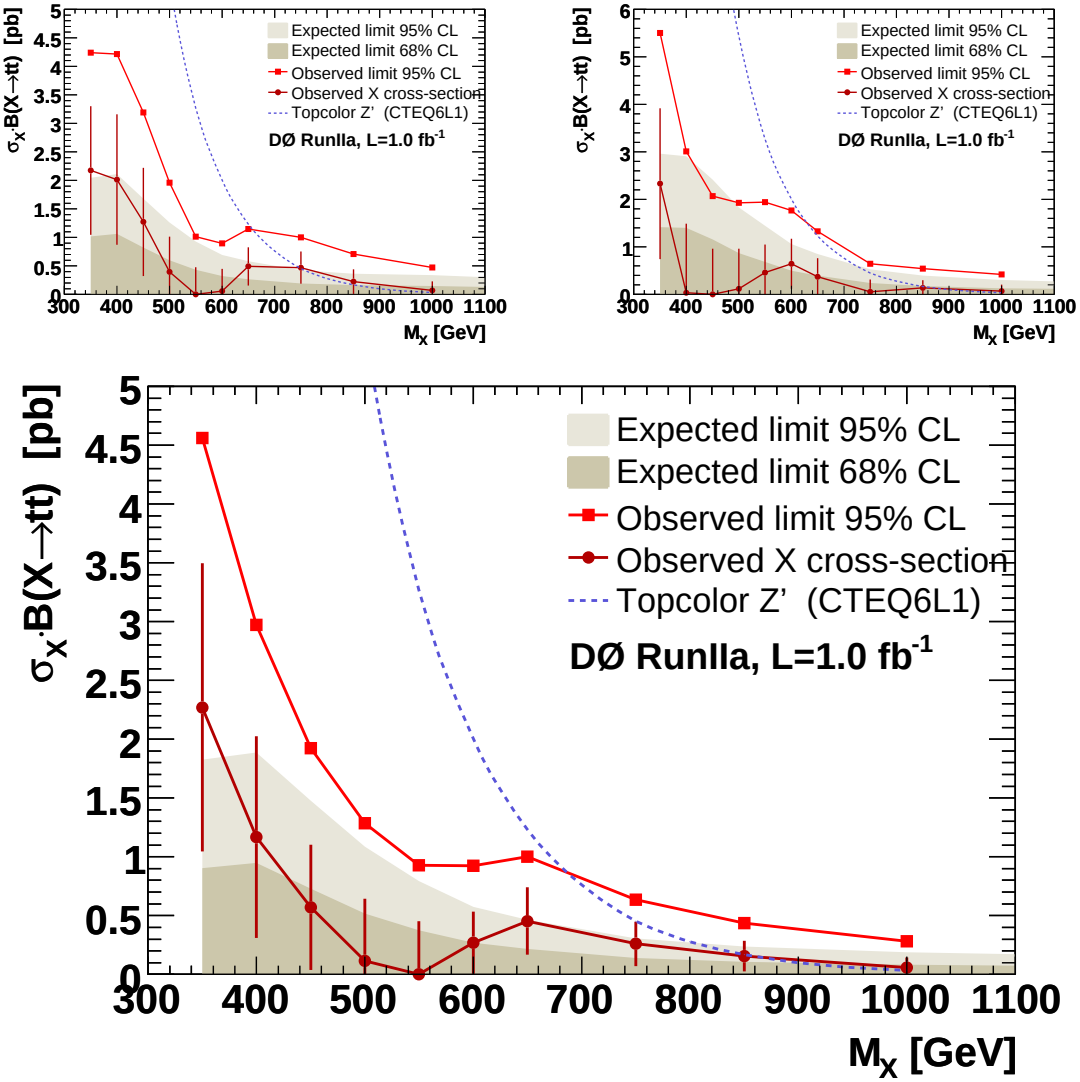


Figure 7.3: The expected 68% and 95% C.L. upper limits on $\sigma_X \times B(X \to t\bar{t})$ (shaded bands) as a function of the resonance mass M_X compared to the observed best resonance cross-section with uncertainty and exclusion limits at the 95% confidence level. The upper plots correspond to individual results from 4-jet (left) and 3-jet (right) bins. The lower plot represents the combined result for the Run IIa dataset.

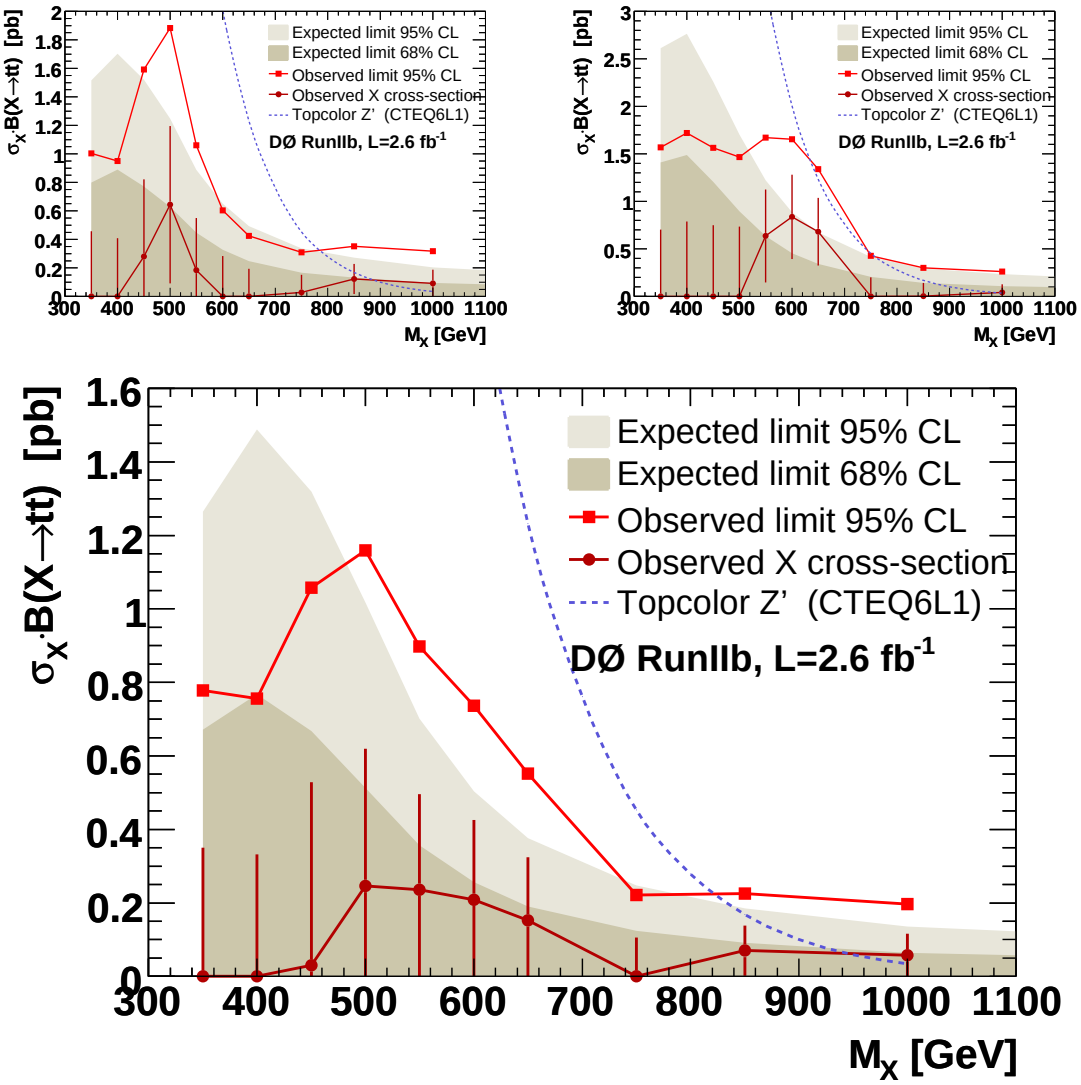


Figure 7.4: The expected 68% and 95% C.L. upper limits on $\sigma_X \times B(X \to t\bar{t})$ (shaded bands) as a function of the resonance mass M_X compared to the observed best resonance cross-section with uncertainty and exclusion limits at the 95% confidence level. The upper plots correspond to individual results from 4-jet (left) and 3-jet (right) bins. The lower plot represents the combined result for the Run IIb dataset.

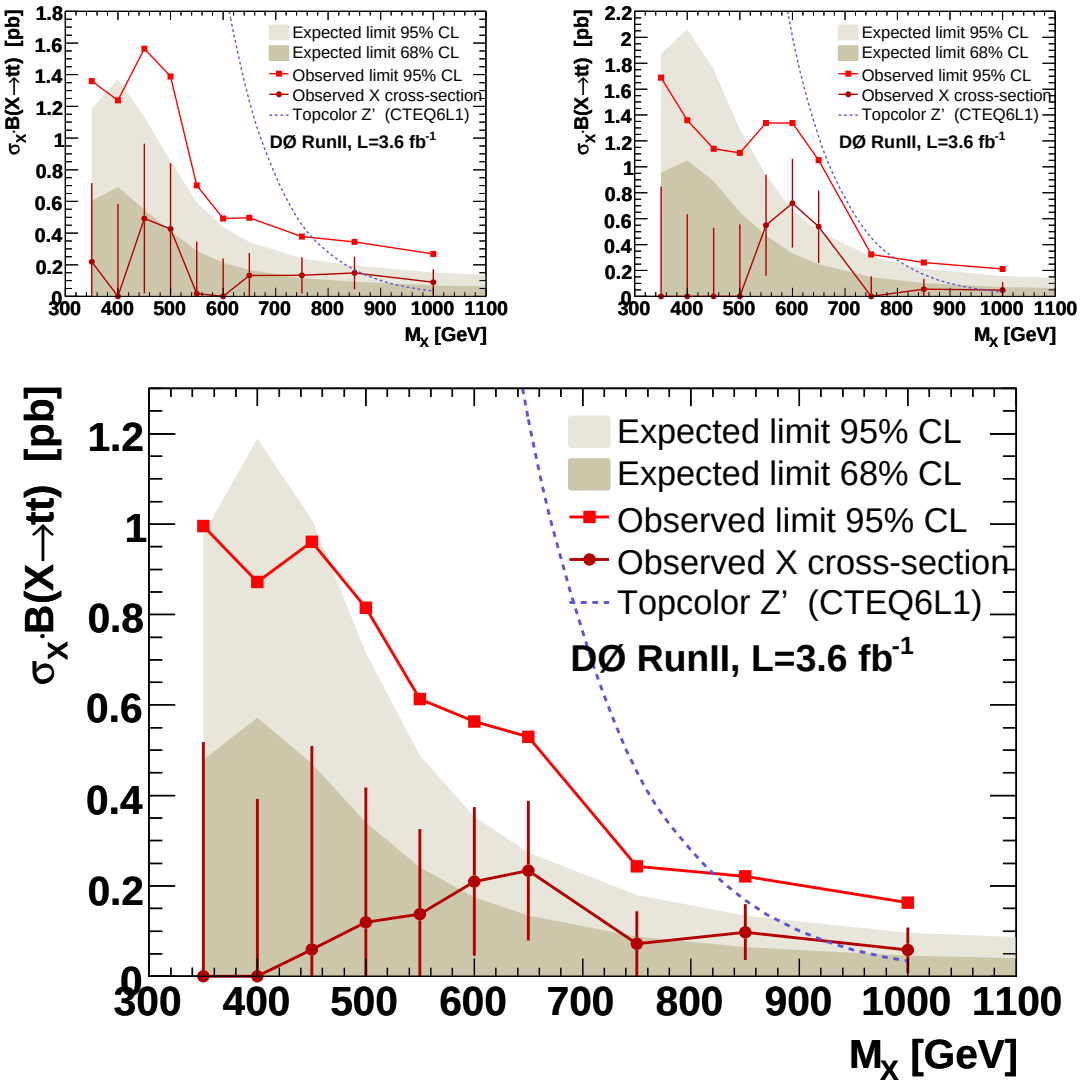


Figure 7.5: The expected 68% and 95% C.L. upper limits on $\sigma_X \times B(X \to t\bar{t})$ (shaded bands) as a function of the resonance mass M_X compared to the observed best resonance cross-section with uncertainty and exclusion limits at the 95% confidence level. The upper plots correspond to individual results from 4-jet (left) and 3-jet (right) bins. The lower plot represents the combined result for the complete dataset.

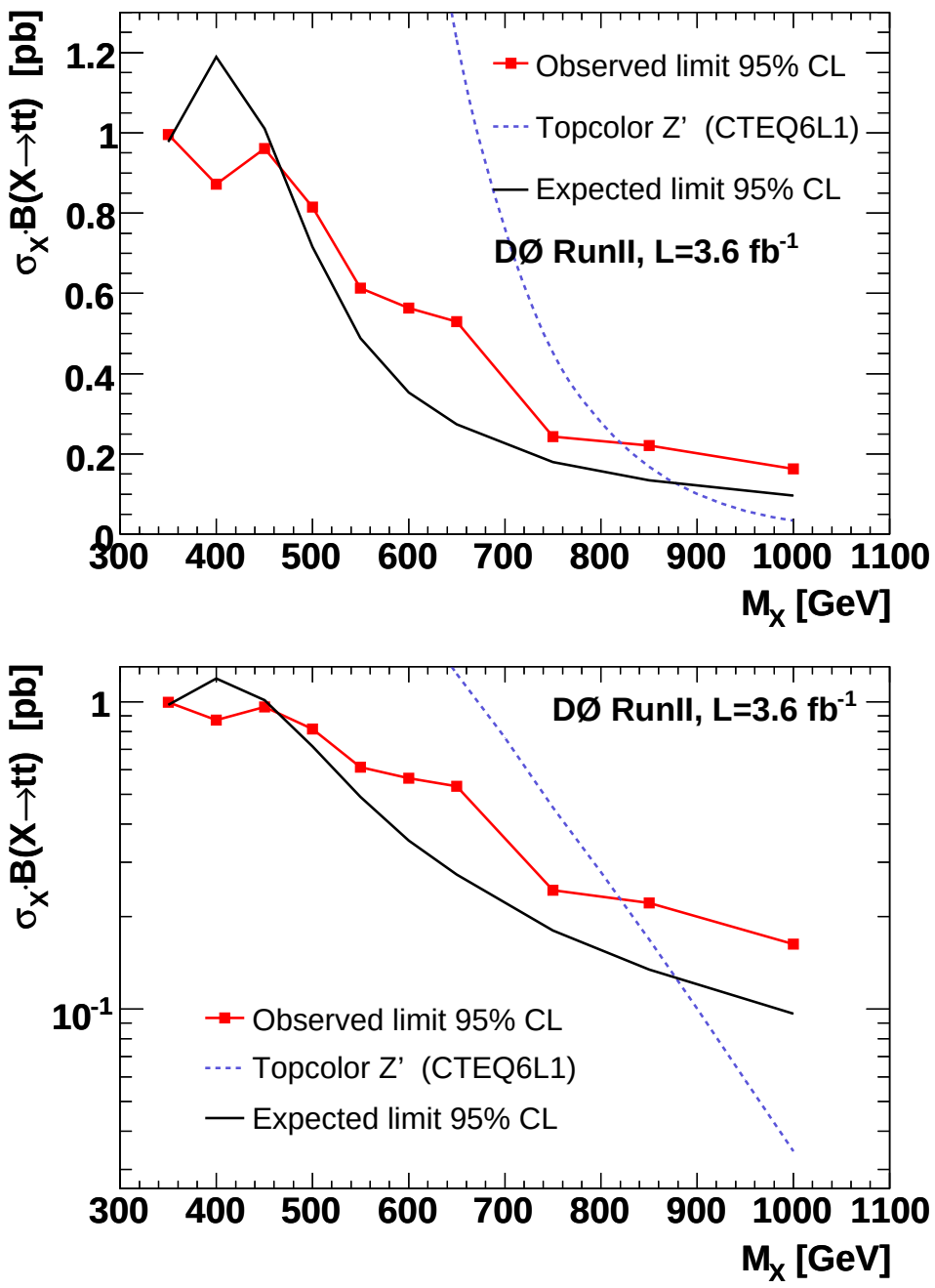


Figure 7.6: Expected and observed $\sigma_X \times B(X \to t\bar{t})$ obtained in $3.6\,\mathrm{fb}^{-1}$.The black curve shows the 95% expected exclusion region, while the red curve shows the corresponding observed 95% C.L. upper limits. The theoretical prediction for a top-color-assisted technicolor Z' boson with a width of $\Gamma_{Z'} = 0.012 M_{Z'}$ as a function of the resonance mass M_X is shown as a dashed line. The upper figure shows the linear plots, while the lower one features a logarithmic cross section axis.

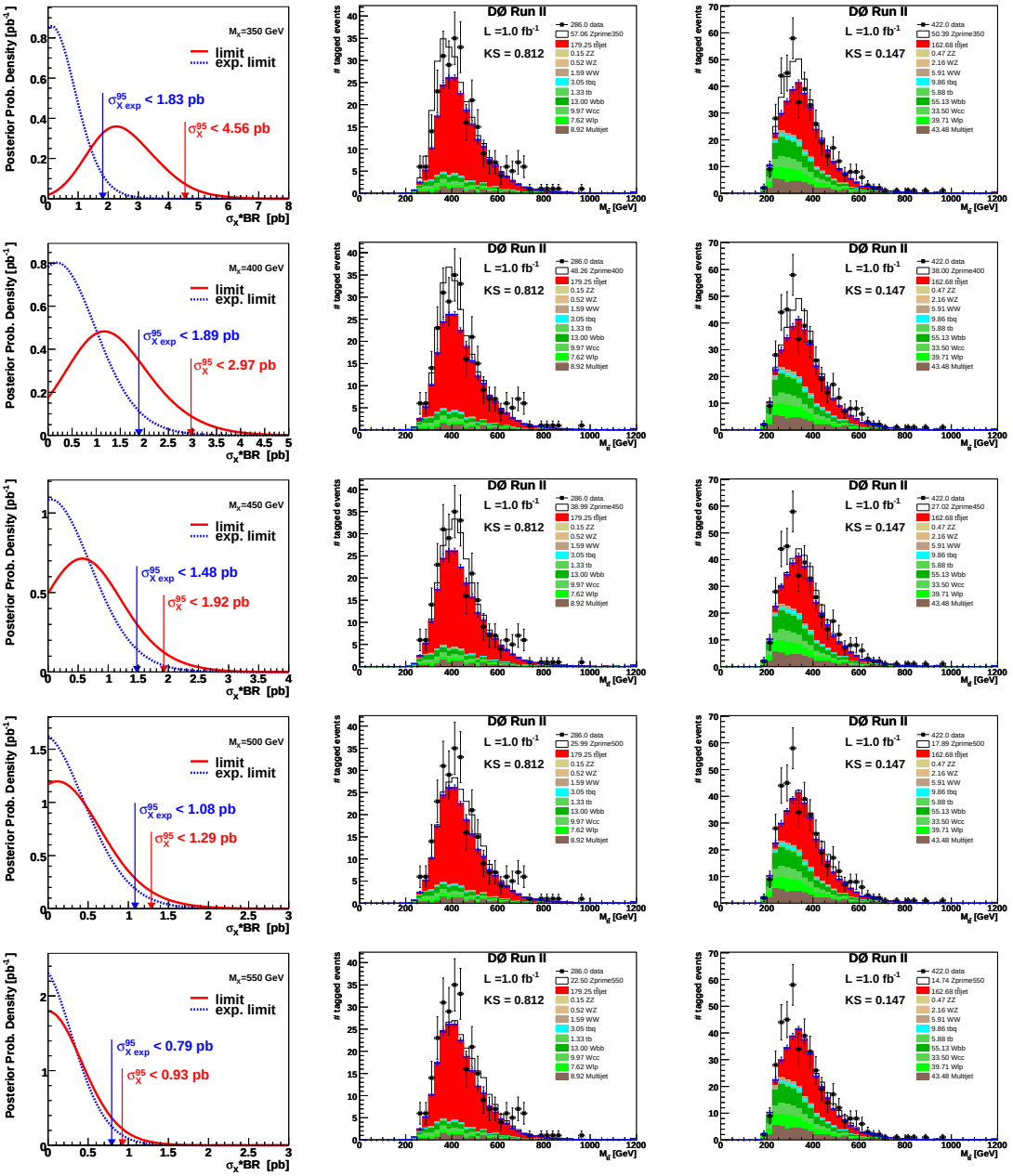


Figure 7.7: In the left plots, the posterior probability densities are shown as a function of $\sigma_X \times B(X \to t\bar{t})$. $M_{t\bar{t}}$ is shown with an assumed resonance contribution in the four or more jet distribution (middle) and the 3-jet distribution (right). Each row corresponds to a different resonance mass assumption, starting from 350 GeV (top) to 550 GeV (bottom) from Run IIa.

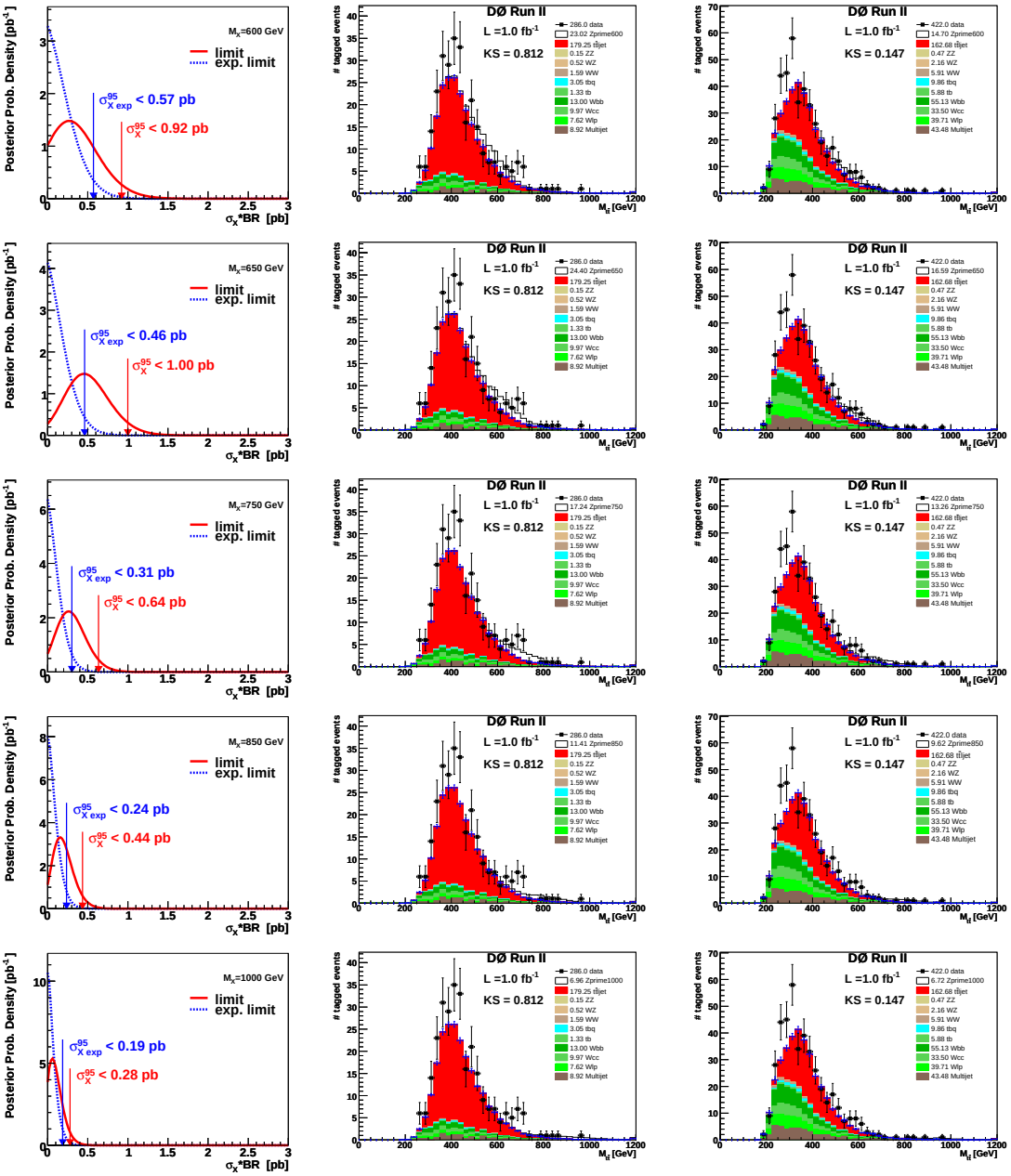


Figure 7.8: In the left plots, the posterior probability densities are shown as a function of $\sigma_X \times B(X \to t\bar{t})$. $M_{t\bar{t}}$ is shown with an assumed resonance contribution in the four or more jet distribution (middle) and the 3-jet distribution (right). Each row corresponds to a different resonance mass assumption, starting from 600 GeV (top) to 1000 GeV (bottom) from Run IIa.

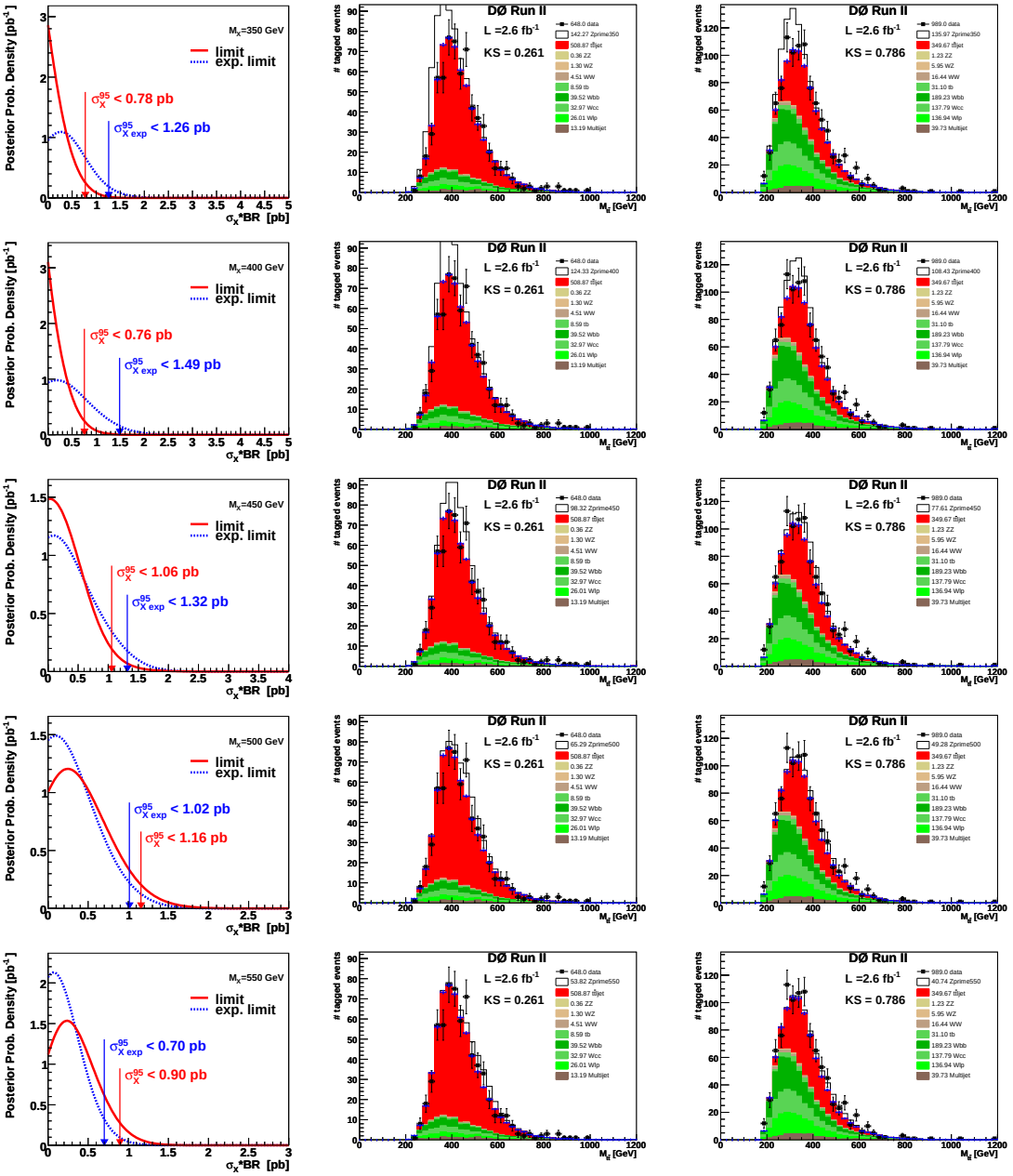


Figure 7.9: In the left plots, the posterior probability densities are shown as a function of $\sigma_X \times B(X \to t\bar{t})$. $M_{t\bar{t}}$ is shown with an assumed resonance contribution in the four or more jet distribution (middle) and the 3-jet distribution (right). Each row corresponds to a different resonance mass assumption, starting from 350 GeV (top) to 550 GeV (bottom) from Run IIb.

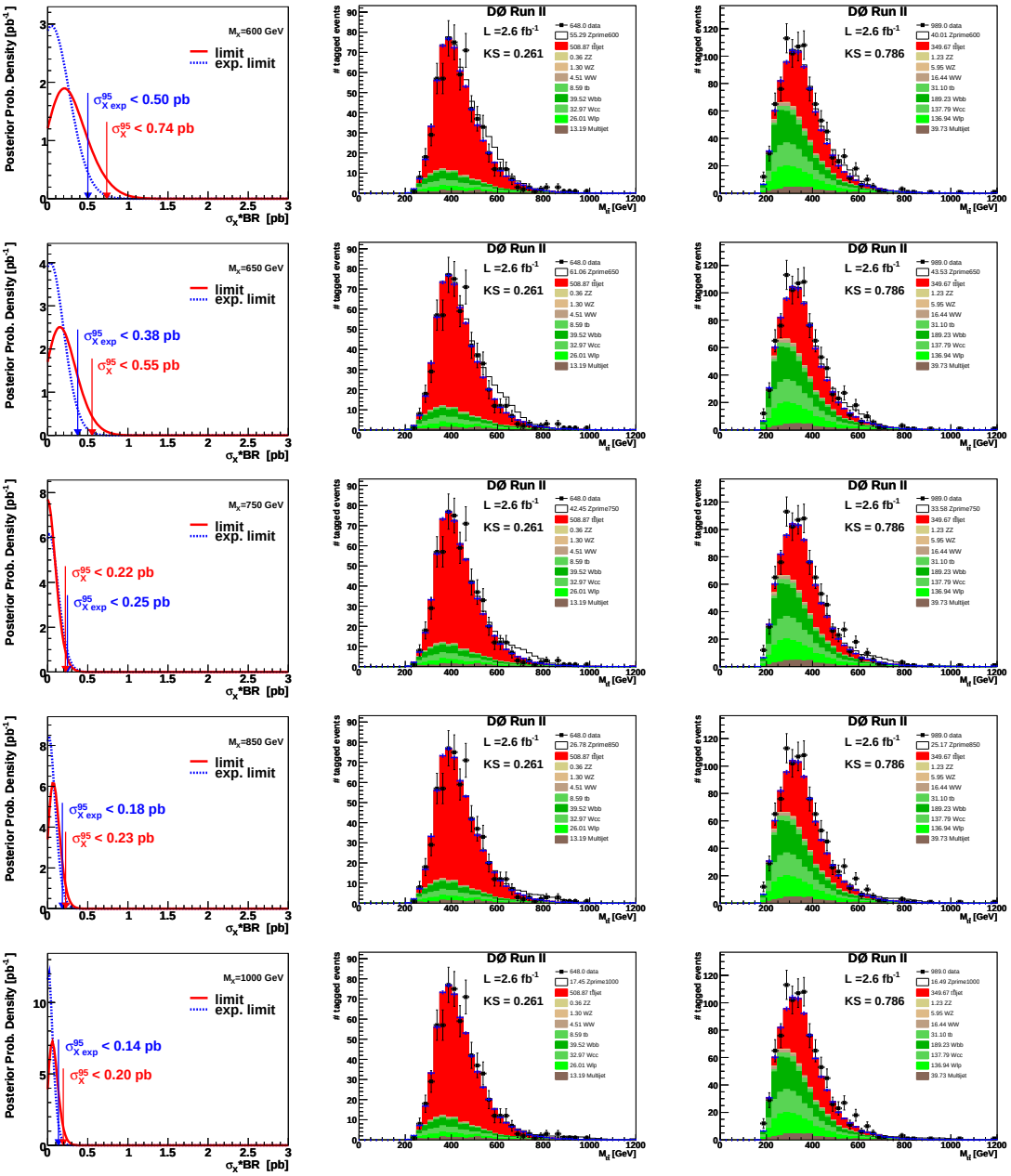


Figure 7.10: In the left plots, the posterior probability densities are shown as a function of $\sigma_X \times B(X \to t\bar{t})$. $M_{t\bar{t}}$ is shown with an assumed resonance contribution in the four or more jet distribution (middle) and the 3-jet distribution (right). Each row corresponds to a different resonance mass assumption, starting from 600 GeV (top) to 1000 GeV (bottom) from Run IIb.

7.3 Ensemble Tests

To check that the procedure is non-biased and gives an estimate on the uncertainty of the limit setting procedure, a set of 5000 pseudo experiments is computed. This procedure is also called ensemble tests. Figure 7.11 show the limits including a yellow band, marking the 68 % contour of the ensemble tests. The expected limit lies well in the middle of the limit and the observed one is mainly inside the yellow band, as expected for one standard deviation. The limit setting procedure is therefore showing no unexpected discrepancies.

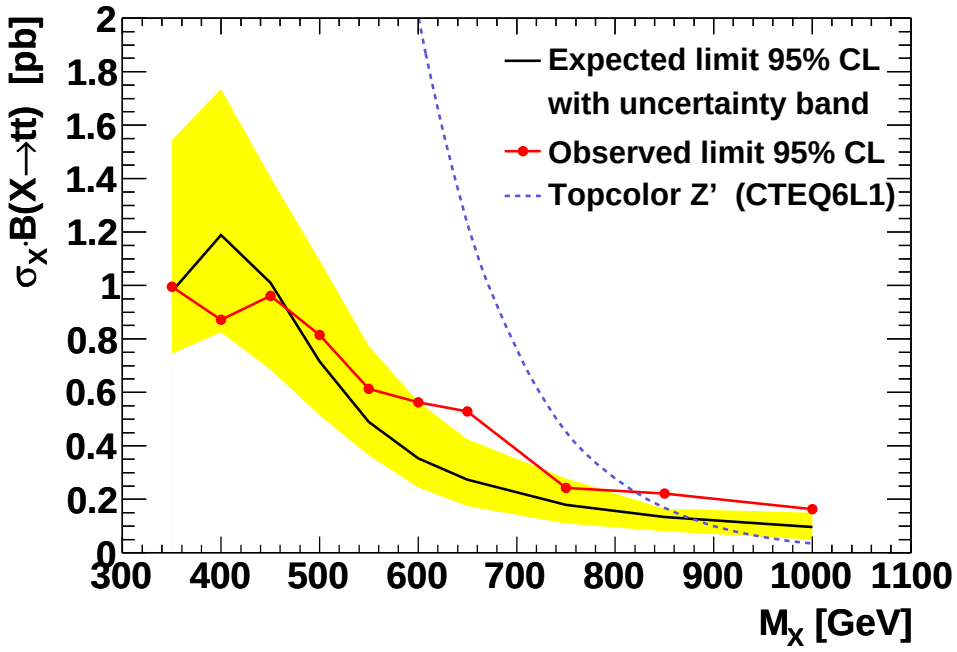


Figure 7.11: The expected and observed 95% confidence level upper limits on $\sigma_X \times B(X \to t\bar{t})$ as a function of the resonance mass M_X for Run II. The shaded band gives the ± 1 sigma uncertainty in the SM expected limit.

7.4 Generalized Couplings

As the Z' model is a benchmark model chosen by CDF and DØ to evaluate a mass limit on a resonance, the result might not be eligible for any other model which has non Z boson like couplings. To broaden the validity of the result resonances with a pure vector or pure axial vector coupling have been studied. The width of the resonance was kept at 1.2 % of the mass as for the Z' resonance. The pure vector and axial vector coupling resonances were studied on the Run IIb dataset with a selected number of resonance masses.

As shown in figure 7.12 the result almost does not depend on the type of the coupling. The change of the result is well below 3 %. As no clear deviation is seen in the examined masses, which range from the low to high end of the mass range, no deviation in the intermediate mass region is to be expected. Therefore results for the limits on the Z' boson can be generalized being model independent for a resonance with small width.

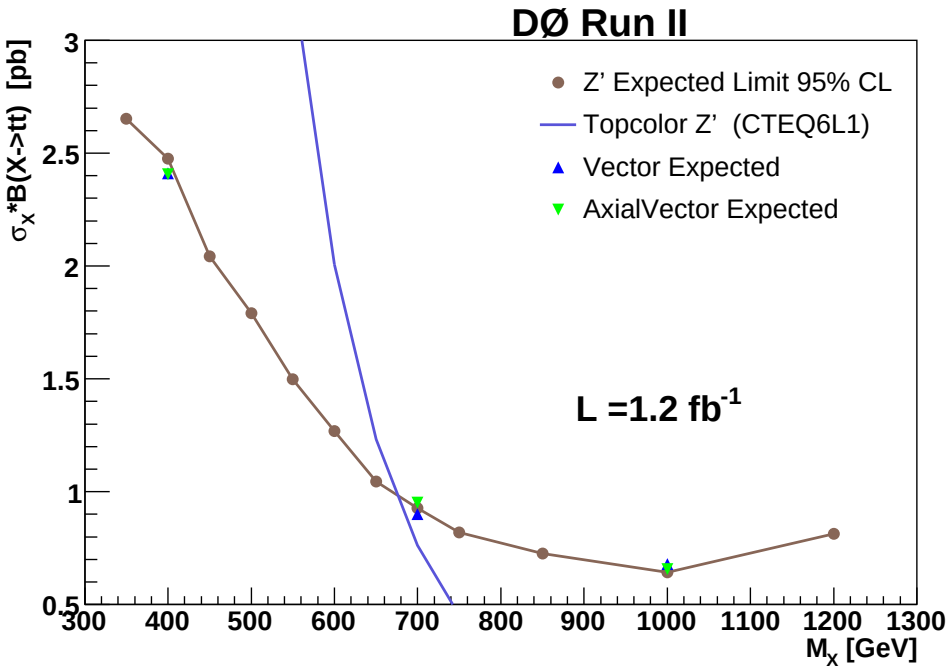


Figure 7.12: Comparison of limits on 95% confidence level upper limits on $\sigma_X \times B(X \rightarrow t\bar{t})$ as a function of resonance mass, comparing Z' type couplings with pure vector or axial vector couplings in three mass points. No dependence on the coupling type can be seen.

The statistical and systematic uncertainty can be judged by the ensemble tests

presented in the last section. The limit values obtained by this study are practically identical, and very far below $1\,\sigma$. The mass resolution of the DØ detector is in the order of 10 %. The analysis method would need further systematic studies, if the resonances width would be larger than this resolution. Migrations between the bins in the analyzed $M_{t\bar{t}}$ distribution would alter the result. This study should therefore be valid for resonances with widths below 5-8 % of their mass. Therefore the limits set on the cross section and mass of the resonance are valid for any small width resonance.

CHAPTER 8

Summary

This analysis presents a search for a narrow width resonance in top quark pair production. The analysis used the l+jets final state. It has been performed using data corresponding to an integrated luminosity of about $3.6\,\mathrm{fb}^{-1}$, collected with the DØ detector during Run II of the Tevatron collider.

After selection cuts, a total of 2345 events survived. The analysis features channels with four or more and the three jets in an event for determining the limit. 1411 events in the 3-jet bin could be analyzed in addition to the 934 events in 4-jet bin. Adding the 3-jet bin to the analysis almost doubled the statistical power of the analysis taking the signal purity into account.

By analyzing the reconstructed $t\bar{t}$ invariant mass distribution, assuming a cross section of 7.48 pb, and using a Bayesian method, model independent upper limits on $\sigma_X \times B(X \rightarrow t\bar{t})$ have been obtained for different hypothesized masses of a narrow-width heavy resonance decaying into $t\bar{t}$. These limits range from 1.0pb at 350 GeV to 0.16pb at 1 TeV.

Within a topcolor-assisted technicolor model, the existence of a leptophobic Z' boson with:

$M_{Z'} < 820$ GeV and width $\Gamma_{Z'} = 0.012 M_{Z'}$ can be excluded at 95% C.L..

CDF and DØ both have made analyses on Z' boson like resonances decaying into $t\bar{t}$ pairs since Run I with increasing sensitivity. Both experiments chose to set the width to 1.2 % of the mass of the resonance. So far the most sensitive result by CDF from 2008 quotes the exclusion of a small width Z' boson to be 805 GeV [190] in the all hadronic channel, while the last l+jets analysis quotes 725 GeV [191]. The more sensitive analysis incorporates about the same number

of systematics as the analysis presented in this thesis. The CDF analysis however, evaluated as well PDF systematics, but did not incorporate it in the final result, as they were found negligible.

The analysis presented in this thesis excludes any Z' resonances with a mass below 820 GeV, which is currently the best exclusion limit for Z' resonances.

The analysis is still statistically limited in the current mass range. The results could be improved to about 880 GeV including the current recorded DØ data of more than 7 fb^{-1}, if the systematical uncertainties would stay the same. With LHC starting to accumulate luminosity fast, the limit can probably be increased there, if the systematic uncertainties are in the same order like at the Tevatron.

Additionally, resonances with pure vector or axial-vector couplings have been examined at certain resonance mass points with the aim of observing the maximal deviation from the Z-like coupling from the benchmark topcolor-assisted technicolor model. Neither of them shows a different behavior concerning the upper limits on $\sigma_X \times B(X \to t\bar{t})$. The mass limit quoted for the benchmark model can therefore be generalized as a small width resonance limit, excluding many other models.

Thus, any small width resonance with a mass below 820 GeVcan be excluded. It is therefore to date the world best exclusion limit for any small width resonance, decaying into $t\bar{t}$ pairs.

APPENDIX A

Determination of the Top Quark Mass

The invariant $t\bar{t}$ mass has further possible applications. Here the extraction of the top quark mass is presented as a method suggested in [45]. Here the rising edge of the differential distribution is used to avoid uncertainties from poorly known high x values for the parton density functions.

A template method is used to extract the top quark mass. In contrast to the previous sections, this analysis uses the 4-jet bin only since the 3-jet bin has a considerable amount of background. This is important as so far no systematic studies have been made to quantify the effect of the background on the mass measurement, in contrast to the resonance analysis.

Templates of the expected signal for various nominal top quark mass values are computed as described in section 6.2 replacing the top quark pair production samples with the ones of different nominal top quark masses. Avoiding an explicit dependency of the top quark cross section can be reached by scaling the events to the total number of observed data. A binned likelihood fit is used to describe the consistency of the various templates with the measured distributions in each channel:

$$\mathcal{L}(m_t) = \sum_{i=1}^{n} \log \frac{(n_i(m_t))^{d_i}}{d_i!} e^{-n_i(m_t)} = \sum_{i=1}^{n} d_i \log n_i(m_t) - n_t(m_t) - \log(d_i!) \quad , \tag{A.1}$$

where $n_i(m_t)$ is the template prediction in the i^{th} bin of the invariant top pair mass distribution for a nominal top quark mass m_t and d_i is the corresponding measurement in observed data. The sum is taken over all bins with the invariant mass $M_{t\bar{t}} < 500\,\text{GeV}$ for which the template from the Monte Carlo simulation predicts a non-zero contribution and at least one data entry exists. This excludes

the bins below 350 GeV and concentrates the fit on the rising edge of the distribution. The log likelihood curve as function of the top mass is fitted with a second order polynomial. The quoted top quark mass corresponds to the maximum of the fitted curve and statistical uncertainty to the mass values for which the log likelihood reduces by 0.5 units. The observed data from the μ+jets for Run IIb has not been used, as part of the samples needed for the mass analysis where unavailable.

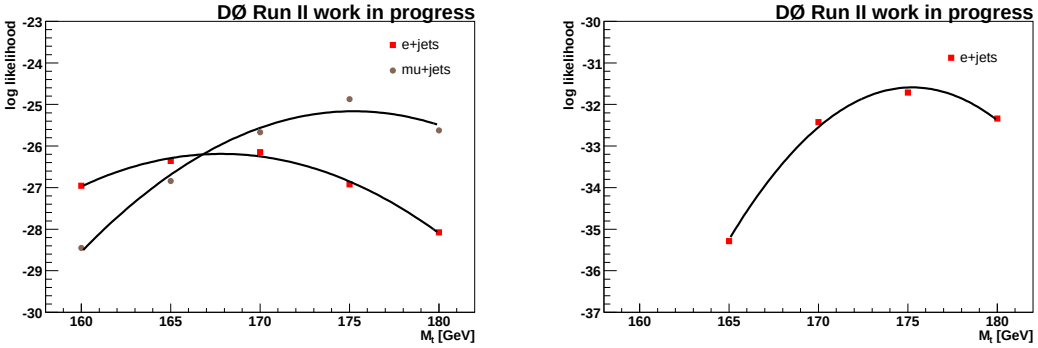


Figure A.1: Observed likelihood curves fitted with a second order polynomial for the individual ℓ+jets channels in Run IIa (left) and Run IIb (right).

The Run IIa data with one or more tags yield the following results:

$$e\text{+jets:} \quad m_t = 167.8 \pm 6.2_{\text{stat}} \tag{A.2}$$

$$\mu\text{+jets:} \quad m_t = 175.2 \pm 5.9_{\text{stat}} \tag{A.3}$$

$$\text{combined:} \quad m_t = 171.8 \pm 4.3_{\text{stat}} \tag{A.4}$$

The Run IIb data with one or more tags yield for the e+jets channel only:

$$e\text{+jets:} \quad m_t = 175.2 \pm 3.8_{\text{stat}} \tag{A.5}$$

This gives an independent method of determining the top quark mass, even measuring a different mass than the $\overline{MS}$ or pole mass. The measurement has about double the statistical uncertainty of the DØ top quark mass measurement in the l+jets channel [192]. The systematic uncertainties would need to be evaluated.

APPENDIX B

Simulation of Top Pair Resonance

Top pair resonance simulation $X \to t\bar{t}$ were produced in DØ MC. For each of the resonance masses individual cardfiles are needed to maintain the desired width of $0.12 M_X$.

For $M_X = 650$ GeV/c^2 and $\Gamma_X = 0.012\ M_X$ the following parameters were used:

- MSTP(51): Set to use the LHPDFCTEQ6L1 parton distribution functions.
- MSTP(52) = 2
- PMAS(6, 1) = 170: This sets the top mass to 170 GeV/c^2.
- MSEL = 21: This selects the process $f_i \bar{f}_i \to \gamma/Z^0/X$.
- MSTP(44) = 3: Only the Z' boson is considered in the matrix elements, no interference terms.
- CKIN(1) = 300, CKIN(2) = 2000, CKIN(3) = 0, CKIN(4) = 2000: The invariant mass of $t\bar{t}$ varies between $300 - 2000\,\mathrm{GeV}/c^2$ and $p_t < 2000.0\,\mathrm{GeV}$.
- CKIN(9) = −5, CKIN(10) = 5, CKIN(11) = −5, CKIN(12) = 5: η ranges for $2 \to 2$ process and for decay products.
- PMAS(32, 1) = 650: Sets the mass of the X boson to 650 GeV/c^2.
- MDME(289, 1) to MDME(293, 1) = 0, MDME(295, 1) to MDME(310, 1) = 0, MDME(294, 1) = 1:
  Only $t\bar{t}$ decay mode is allowed for the X boson.
- Vector and axial couplings of the X boson to first generation quark and leptons:

PARU(121) = −0.418, PARU(122) = −0.603,
PARU(123) = 0.233, PARU(124) = 0.603,
PARU(125) = −0.048, PARU(126) = −0.603,
PARU(127) = 0.603, PARU(128) = 0.603

- Vector and axial couplings of the X boson to second generation quark and leptons:
  PARJ(180) = −0.418, PARJ(181) = −0.603,
  PARJ(182) = 0.233, PARJ(183) = 0.603,
  PARJ(184) = −0.048, PARJ(185) = −0.603,
  PARJ(186) = 0.603, PARJ(187) = 0.603
- Vector and axial couplings of the X boson to third generation quark and leptons:
  PARJ(188) = −0.418, PARJ(189) = −0.603,
  PARJ(190) = 0.233, PARJ(191) = 0.603,
  PARJ(192) = −0.048, PARJ(193) = −0.603,
  PARJ(194) = 0.603, PARJ(195) = 0.603

The settings for the vector and axial couplings are needed to set the width of the X boson to the desired value of $\Gamma_X = 0.012 M_X$ since the direct way of setting the width in Pythia does not work because one needs to include the full interference structure. In Pythia, the default setting for the couplings of the X boson to quarks and leptons is the same as that for the Standard Model Z boson.

APPENDIX C

Control-Plots

As a sanity check the following control plots are shown:

- Run IIa
  - Figures C.1 - C.6 show data-MC comparisons after all preselection cuts except for the b-tag.
  - Figures C.7 - C.9 show data-MC comparisons of 3-jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.10 - C.12 show data-MC comparisons of 3-jet events after all preselection cuts including two or more b-tags using the Medium working point of the NN tagger.
  - Figures C.13 - C.18 show data-MC comparisons of four or more jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.16 - C.18 show data-MC comparisons of four or more jet events after all preselection cuts including two or more b-tags using the Medium working point of the NN tagger.
  - Figure C.19 shows the $t\bar{t}$ invariant mass distributions for three or more jet events in the e+jets and μ+jets channels.
  - Figure C.20 shows the $t\bar{t}$ invariant mass distributions for four or more jet events in the e+jets and μ+jets channels.
  - Figures C.21 - C.26 show data-MC comparisons after all preselection cuts except for the b-tag for one jet events.
  - Figures C.24 - C.26 show data-MC comparisons after all preselection cuts except for the b-tag for two jet events.

  - Figures C.27 - C.29 show data-MC comparisons of 1-jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.30 - C.32 show data-MC comparisons of 2-jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.33 - C.35 show data-MC comparisons of 2-jet events after all preselection cuts including two or more b-tags using the Medium working point of the NN tagger.
- Run IIb
  - Figures C.36 - C.41 show data-MC comparisons after all preselection cuts except for the b-tag.
  - Figures C.42 - C.44 show data-MC comparisons of 3-jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.45 - C.47 show data-MC comparisons of 3-jet events after all preselection cuts including two or more b-tags using the Medium working point of the NN tagger.
  - Figures C.48 - C.53 show data-MC comparisons of four or more jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.51 - C.53 show data-MC comparisons of four or more jet events after all preselection cuts including two or more b-tags using the Medium working point of the NN tagger.
  - Figure C.54 shows the $t\bar{t}$ invariant mass distributions for three or more jet events in the e+jets and μ+jets channels.
  - Figure C.55 shows the $t\bar{t}$ invariant mass distributions for four or more jet events in the e+jets and μ+jets channels.
  - Figures C.56 - C.61 show data-MC comparisons after all preselection cuts except for the b-tag for 1-jet events.
  - Figures C.59 - C.61 show data-MC comparisons after all preselection cuts except for the b-tag for 2-jet events.
  - Figures C.62 - C.64 show data-MC comparisons of 1-jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.65 - C.67 show data-MC comparisons of 2-jet events after all preselection cuts including one b-tag using the Medium working point of the NN tagger.
  - Figures C.68 - C.70 show data-MC comparisons of 2-jet events after all preselection cuts including two or more b-tags using the Medium working point of the NN tagger.

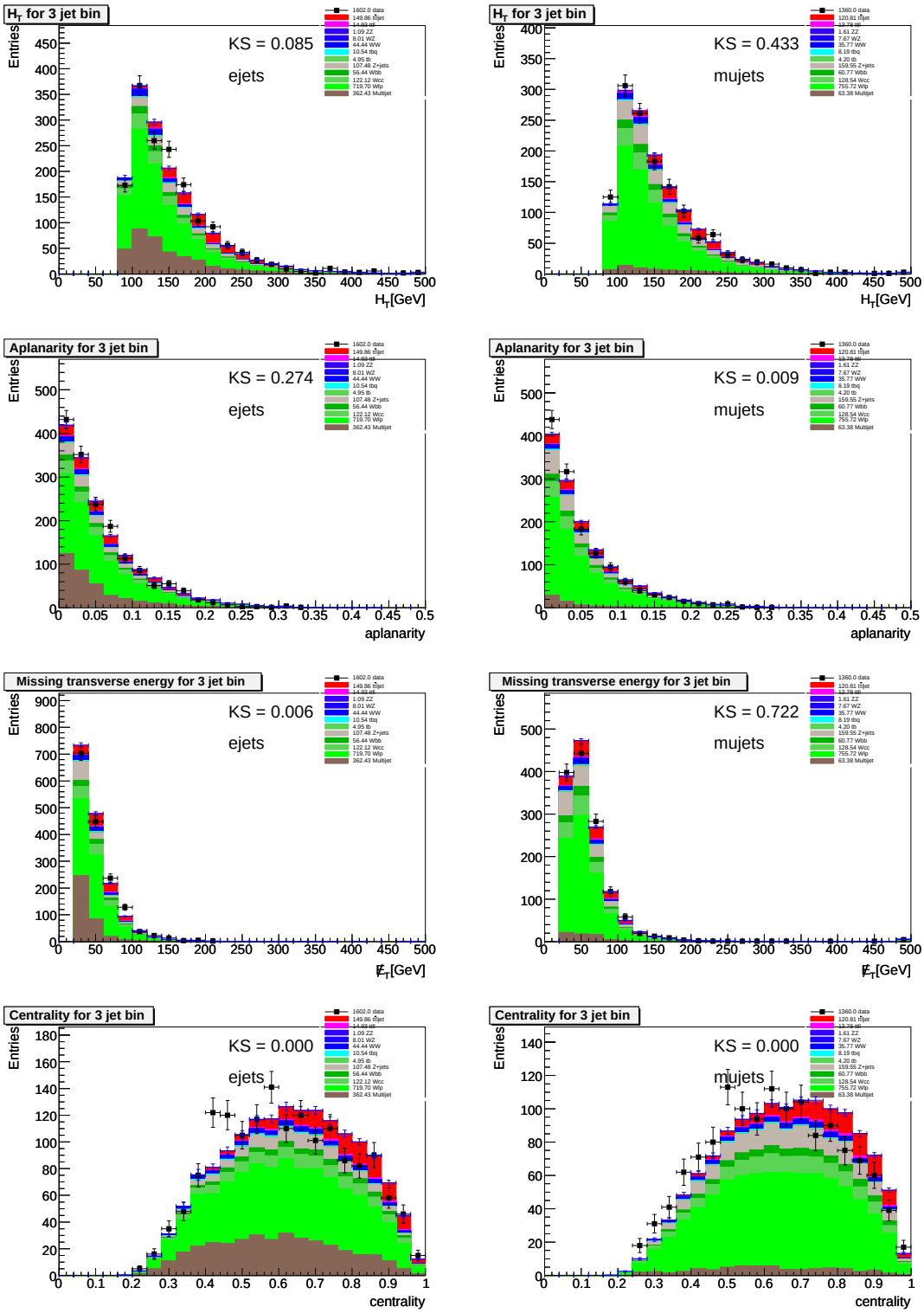


Figure C.1: Comparison for Run IIa between data and Monte Carlo with three jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

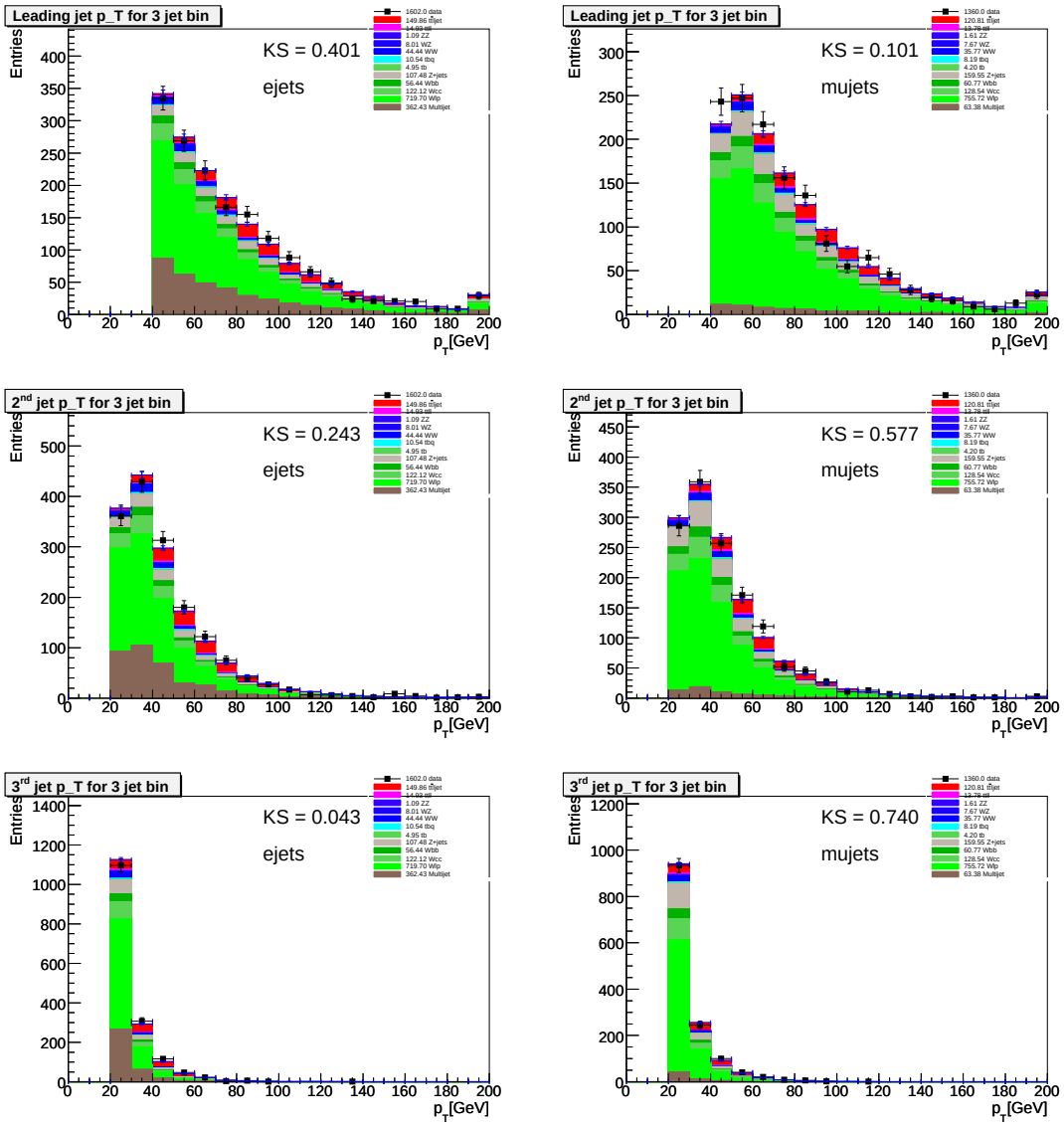


Figure C.2: Comparison for Run IIa between data and Monte Carlo with three jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

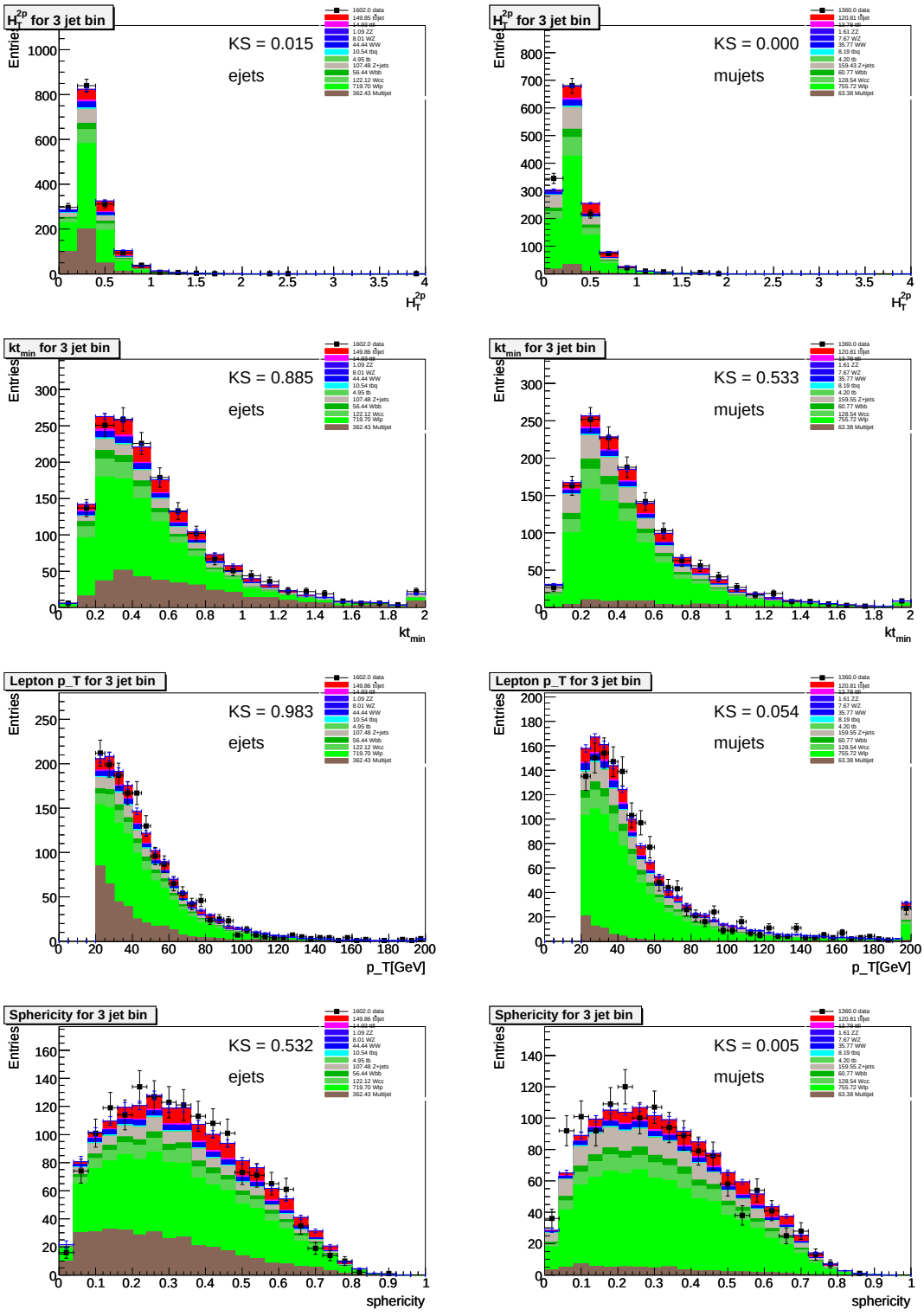


Figure C.3: Comparison for Run IIa between data and Monte Carlo with three jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

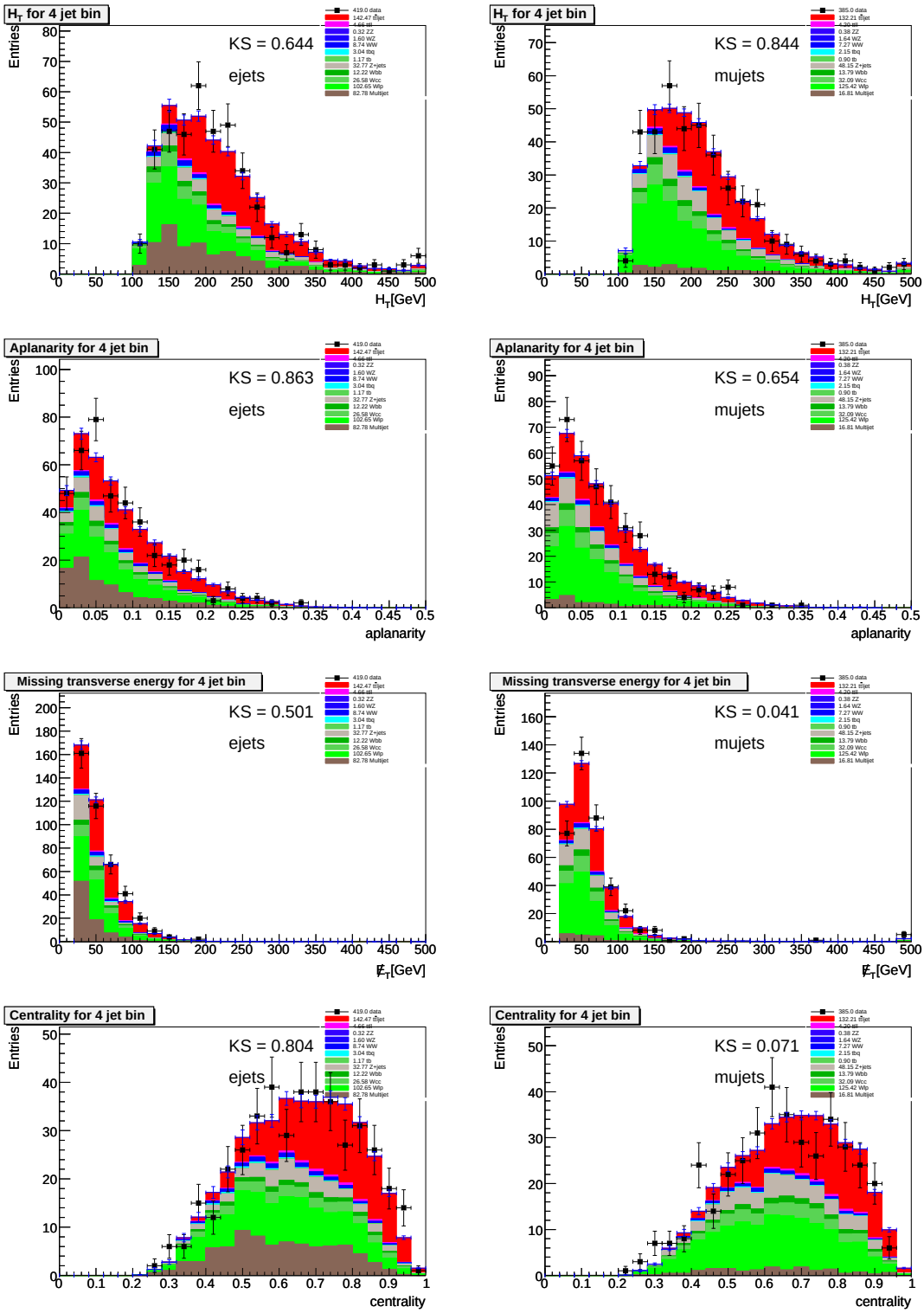


Figure C.4: Comparison for Run IIa between data and Monte Carlo with four or more jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

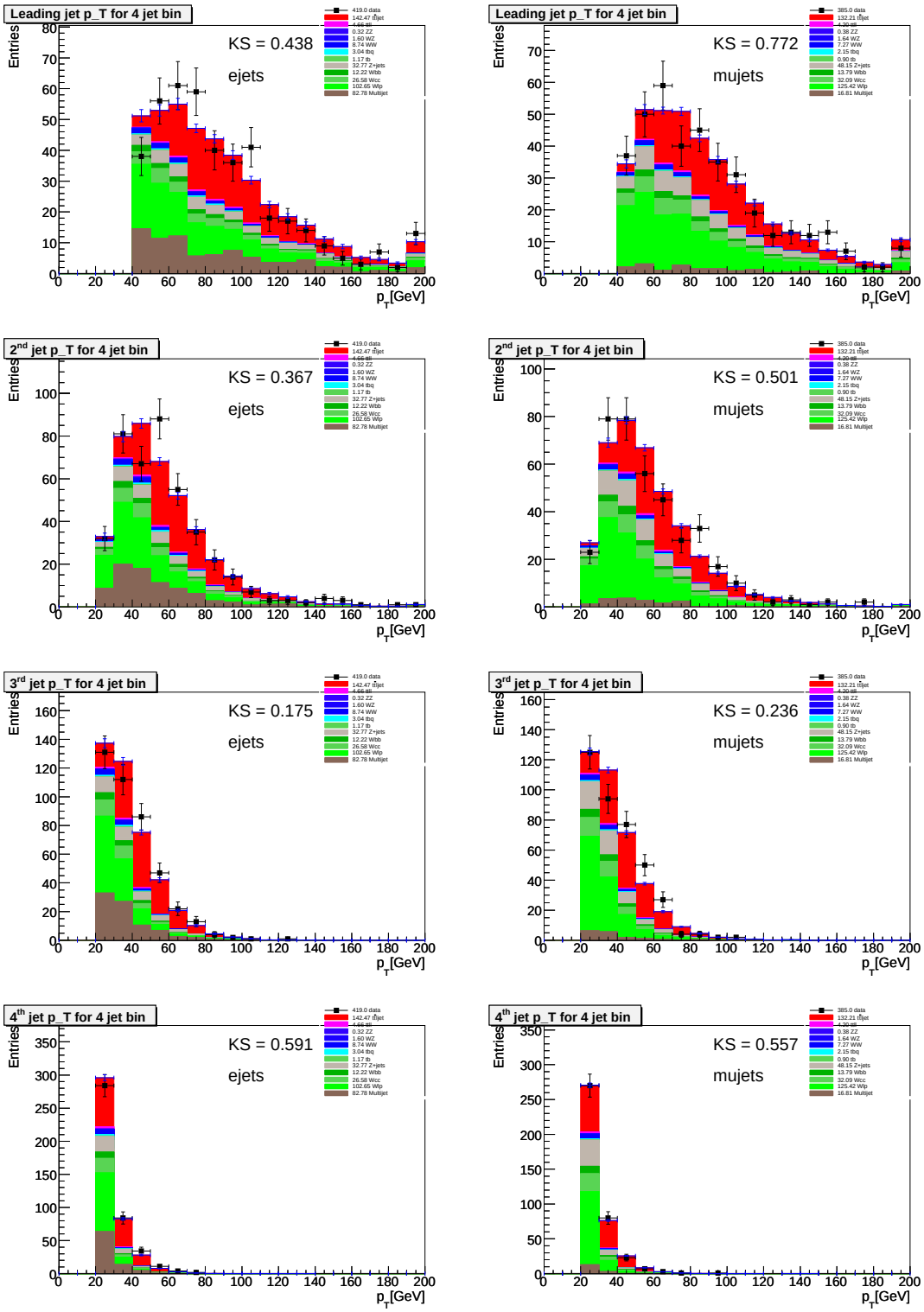


Figure C.5: Comparison for Run IIa between data and Monte Carlo with four or more jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

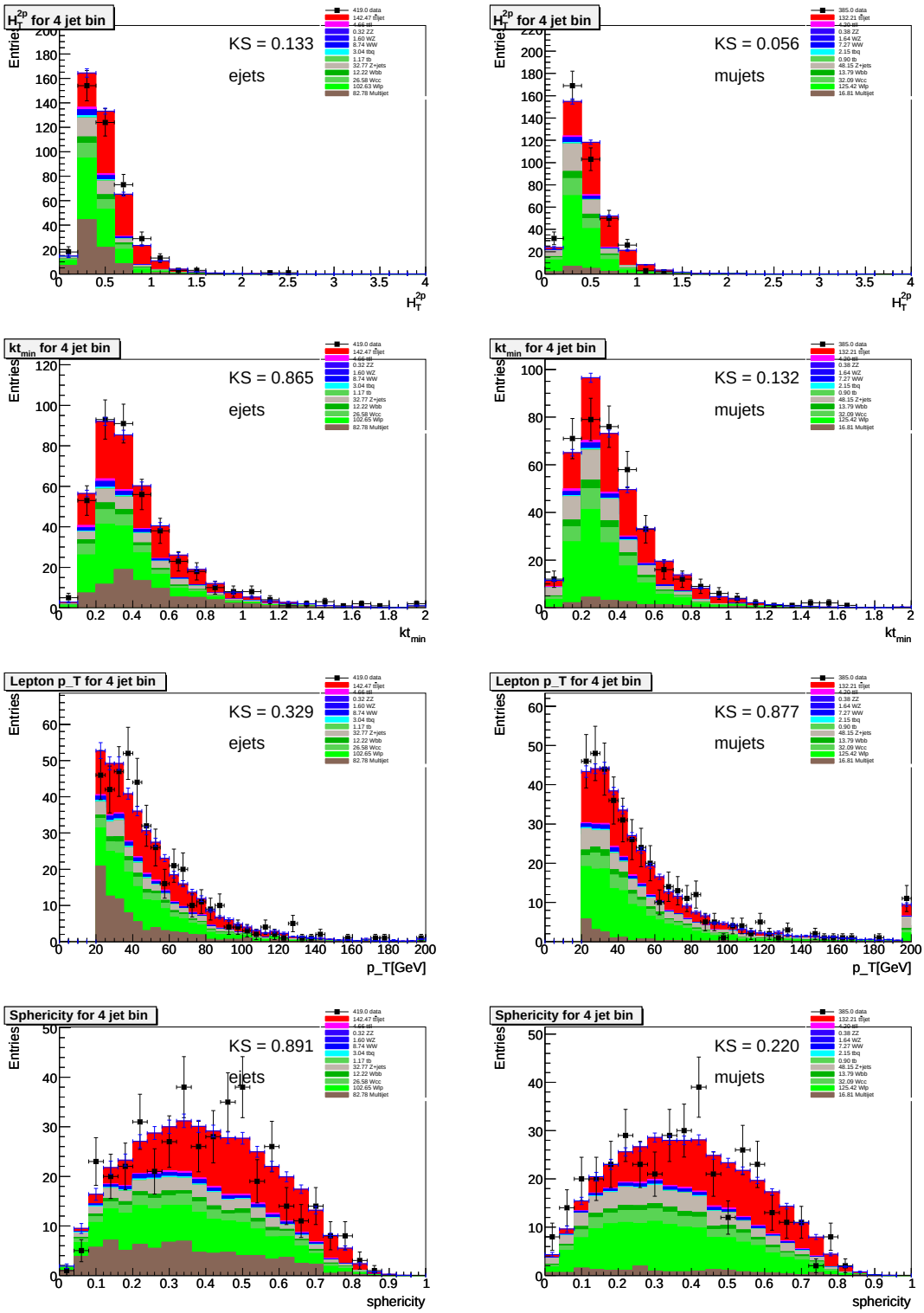


Figure C.6: Comparison for Run IIa between data and Monte Carlo with four or more jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

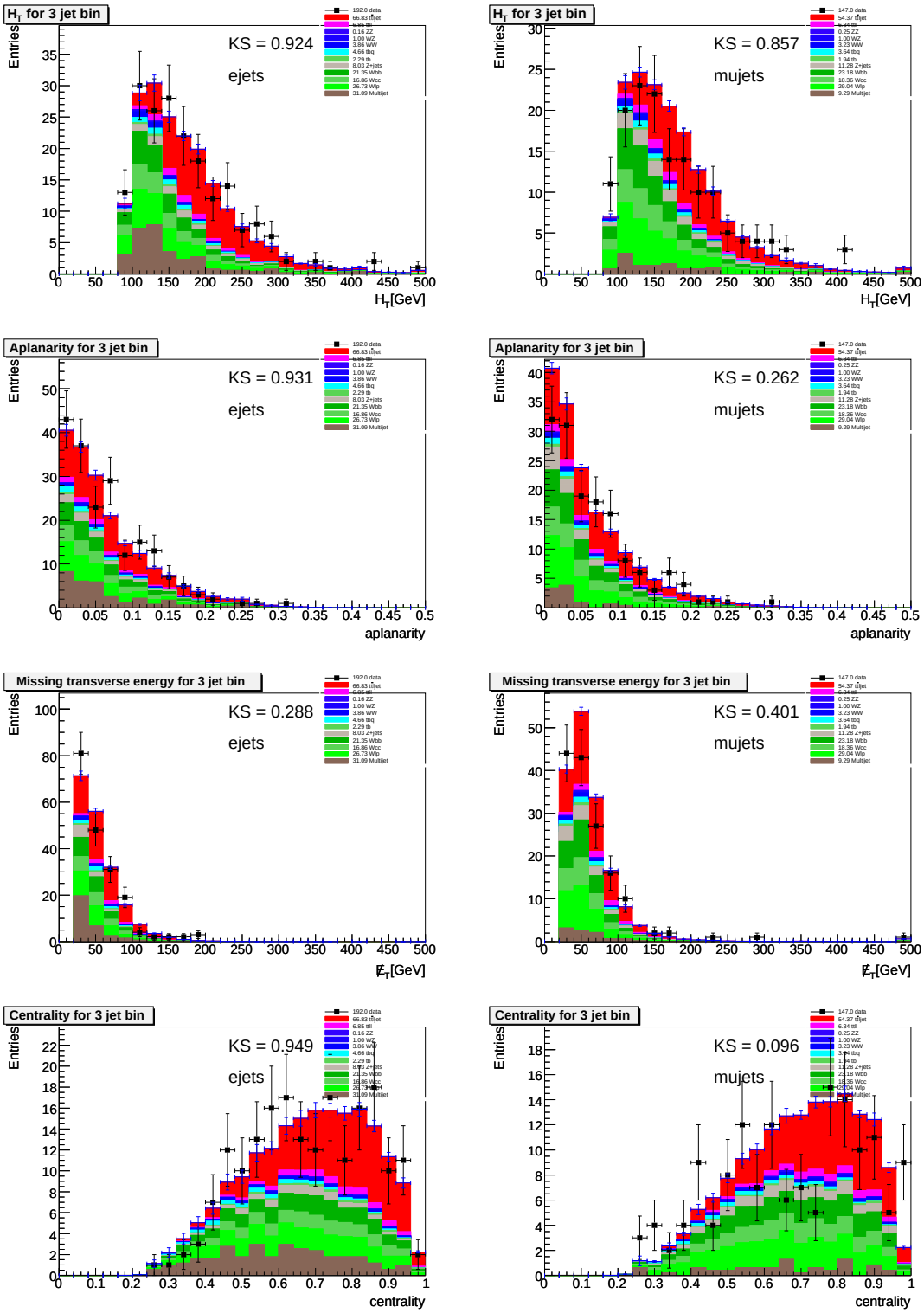


Figure C.7: Comparison for Run IIa between data and Monte Carlo for three jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

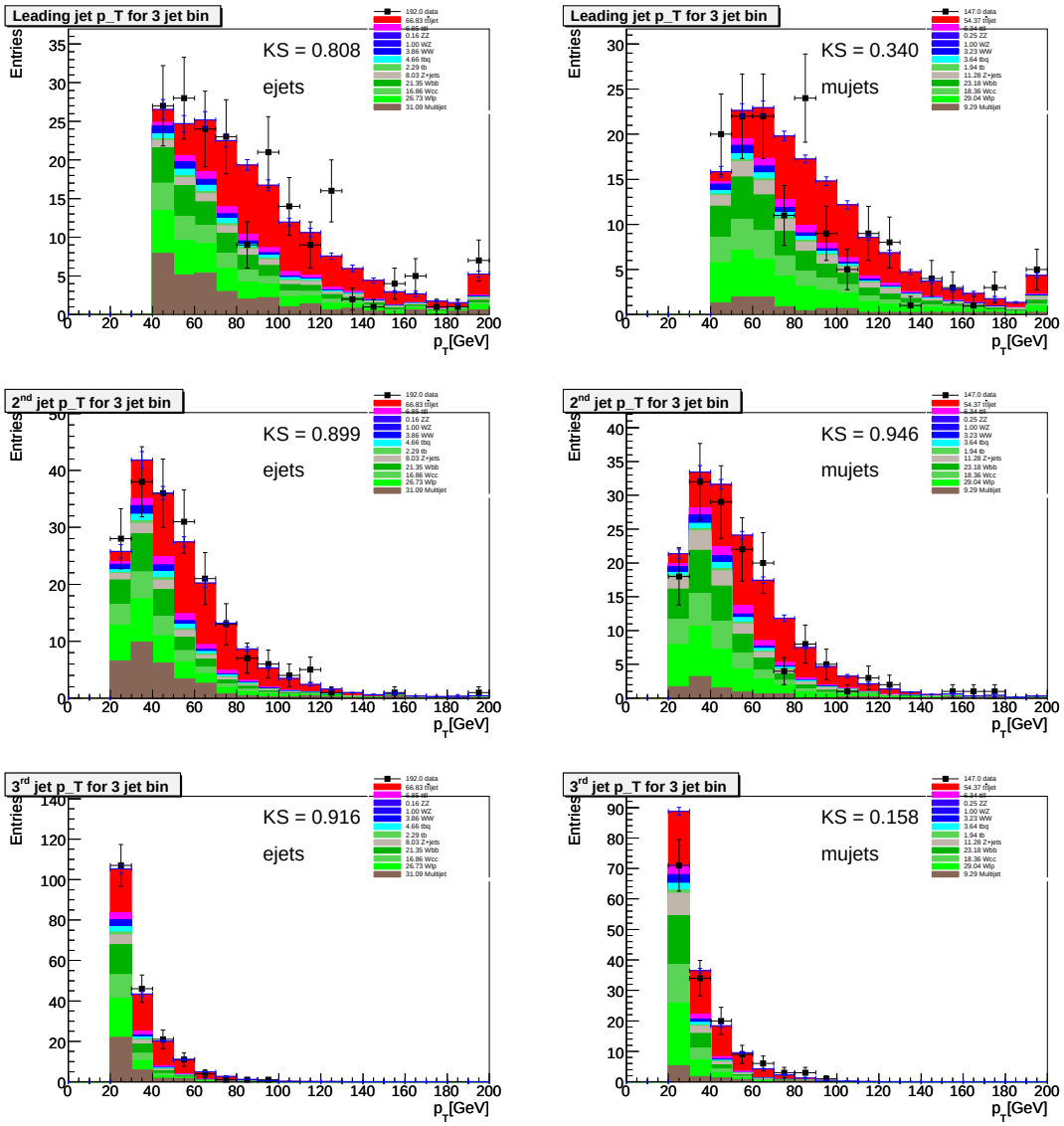


Figure C.8: Comparison for Run IIa between data and Monte Carlo for three jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet and p_T of the third leading jet.

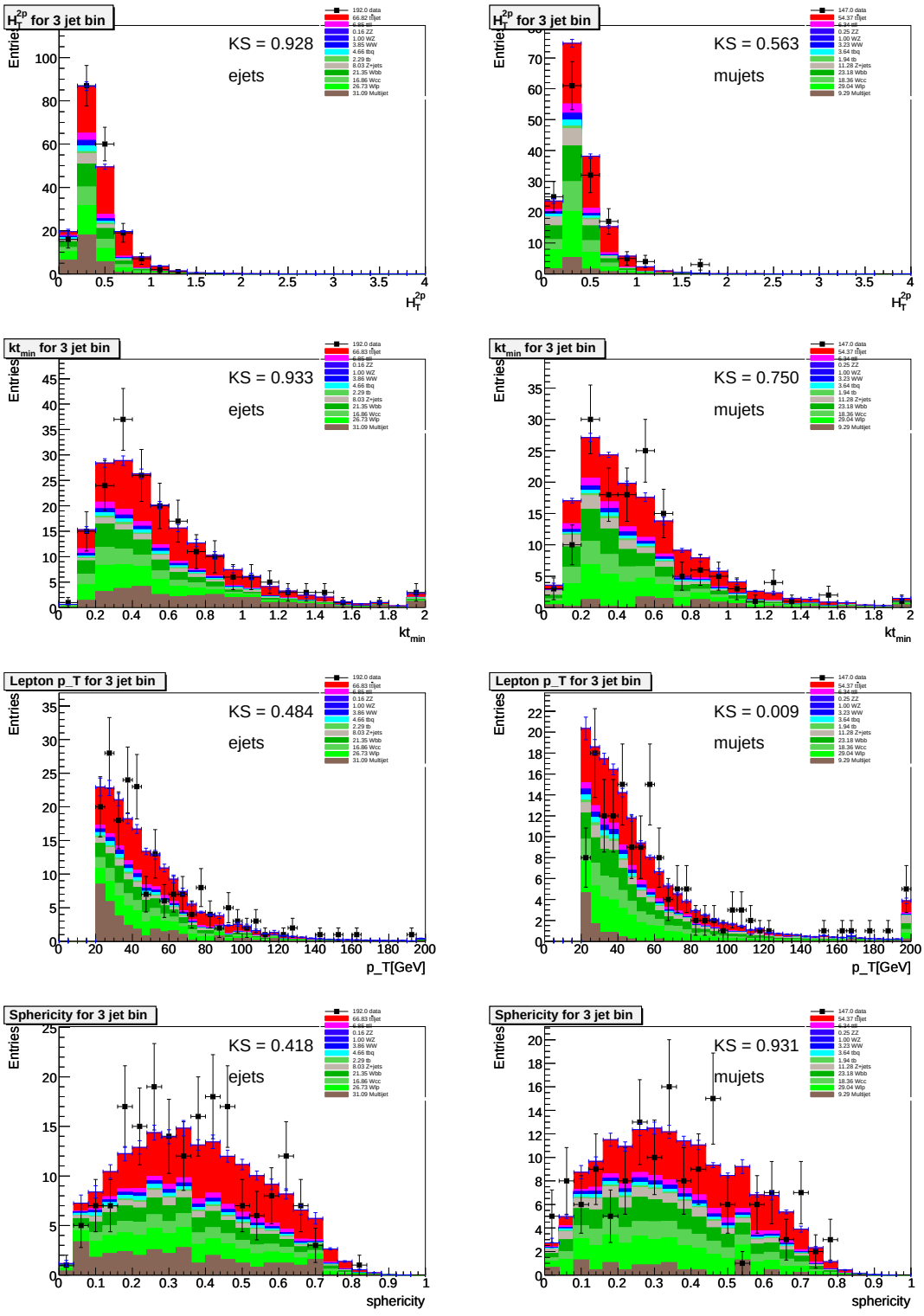


Figure C.9: Comparison for Run IIa between data and Monte Carlo for three jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

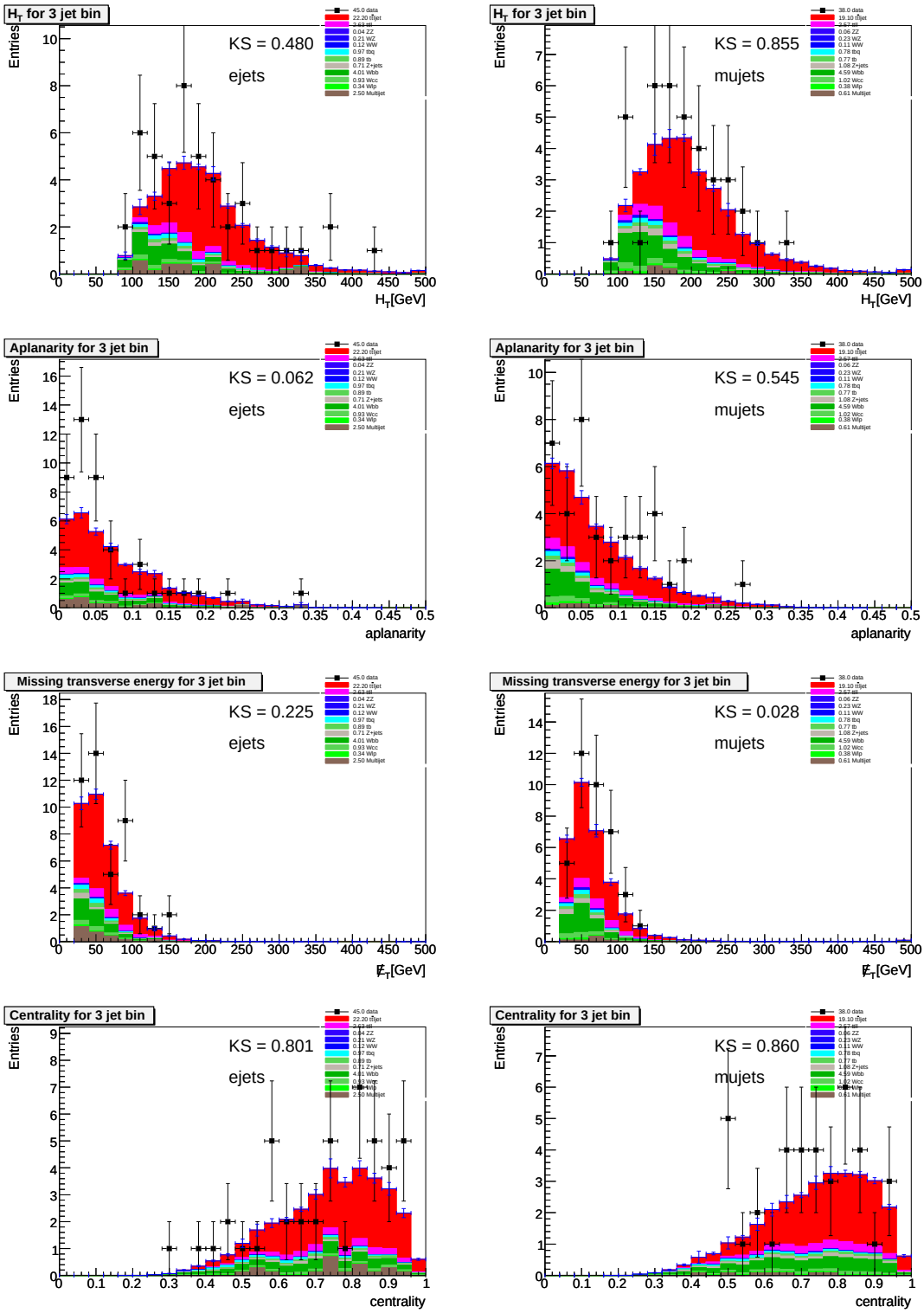


Figure C.10: Comparison for Run IIa between data and Monte Carlo for three jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

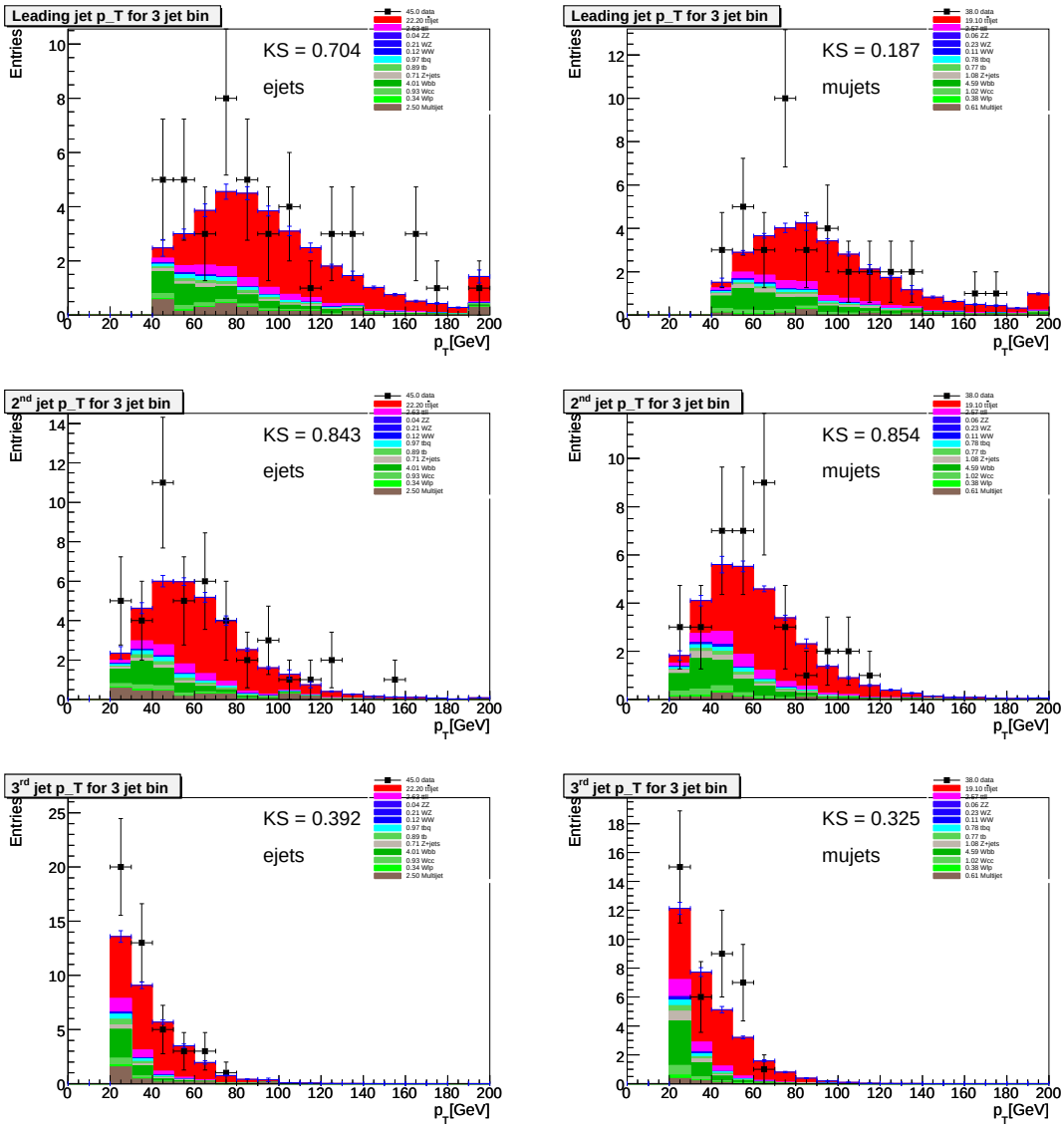


Figure C.11: Comparison for Run IIa between data and Monte Carlo for three jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet and p_T of the third leading jet.

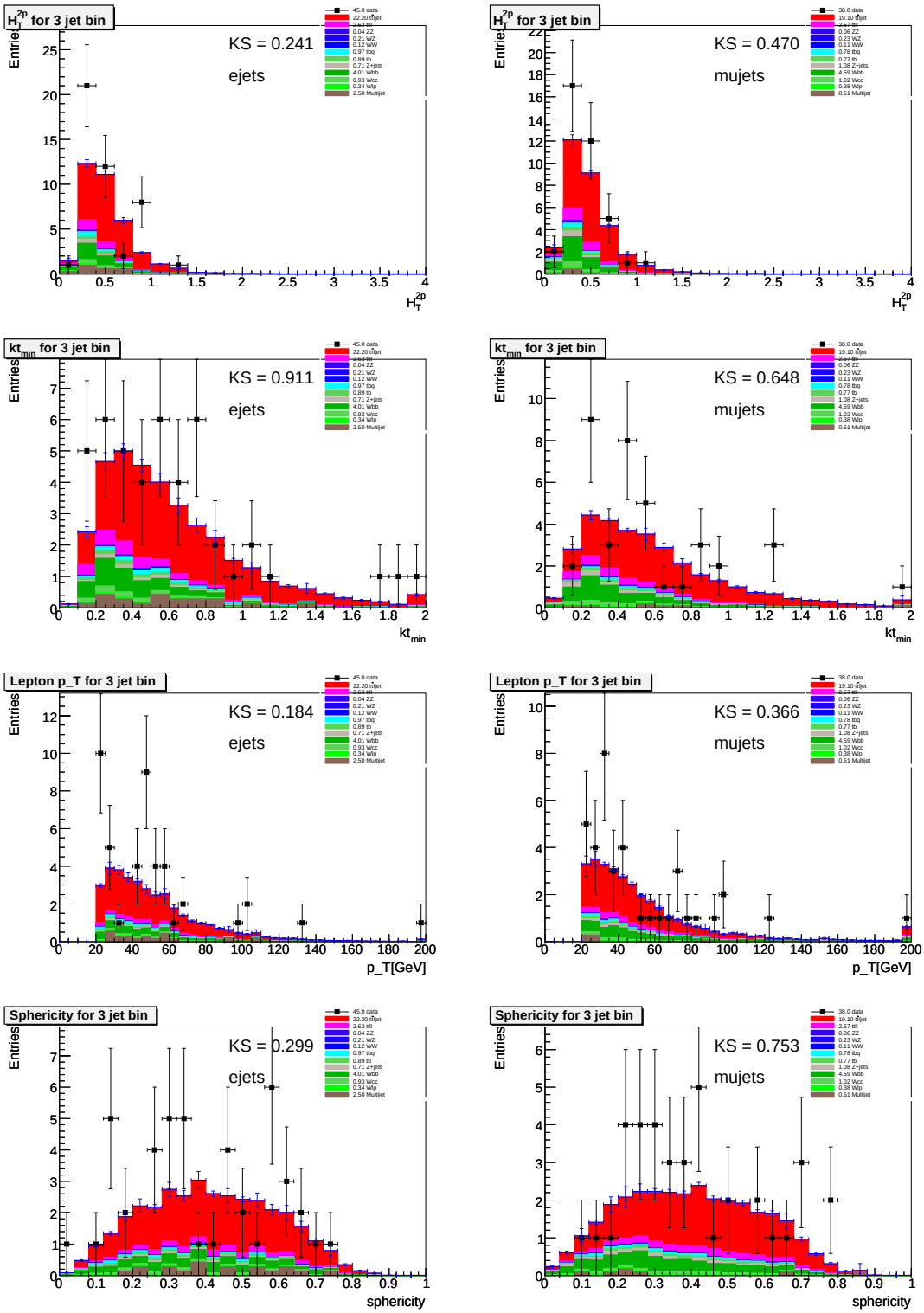


Figure C.12: Comparison for Run IIa between data and Monte Carlo for three jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

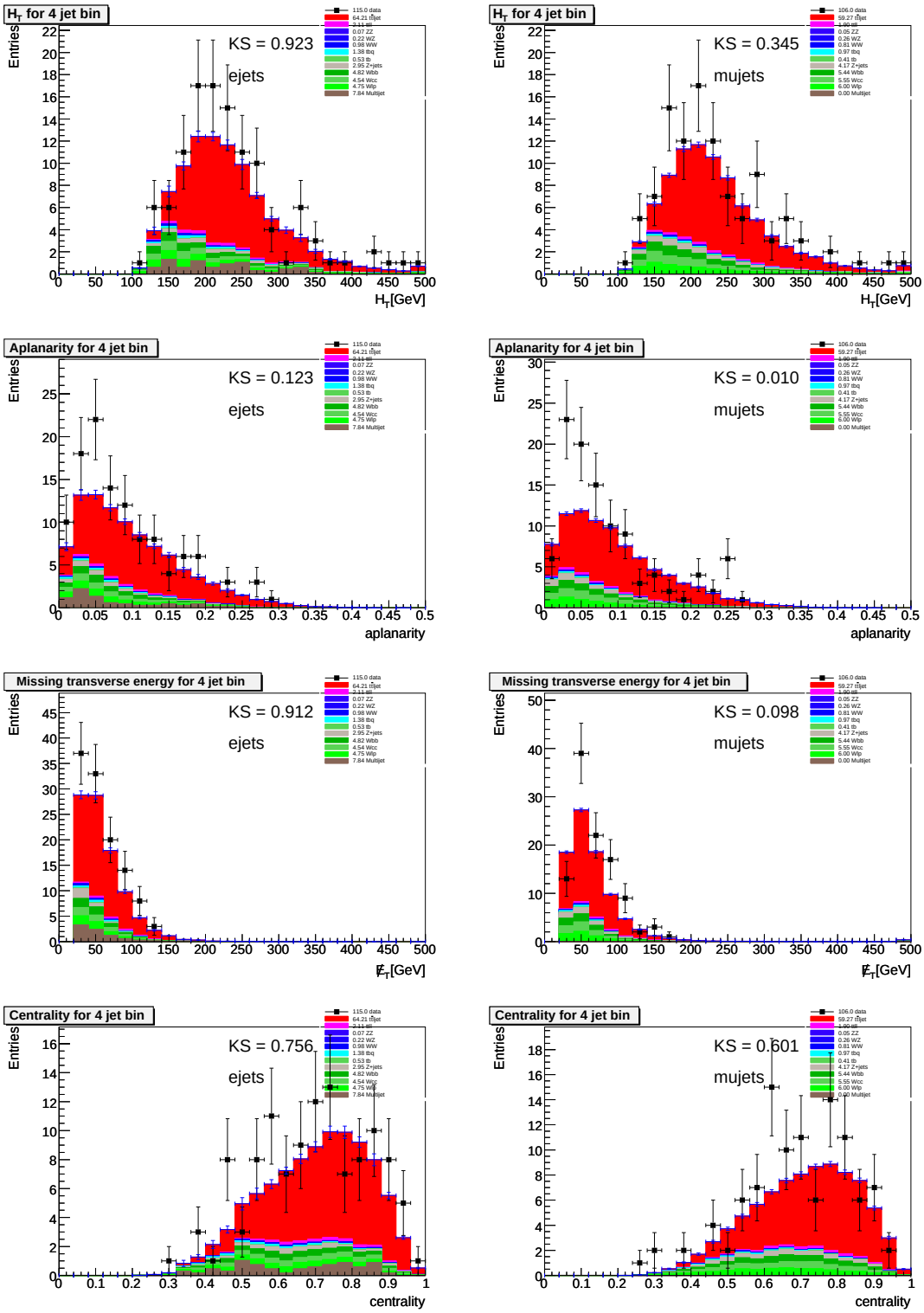


Figure C.13: Comparison for Run IIa between data and Monte Carlo for four or more jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

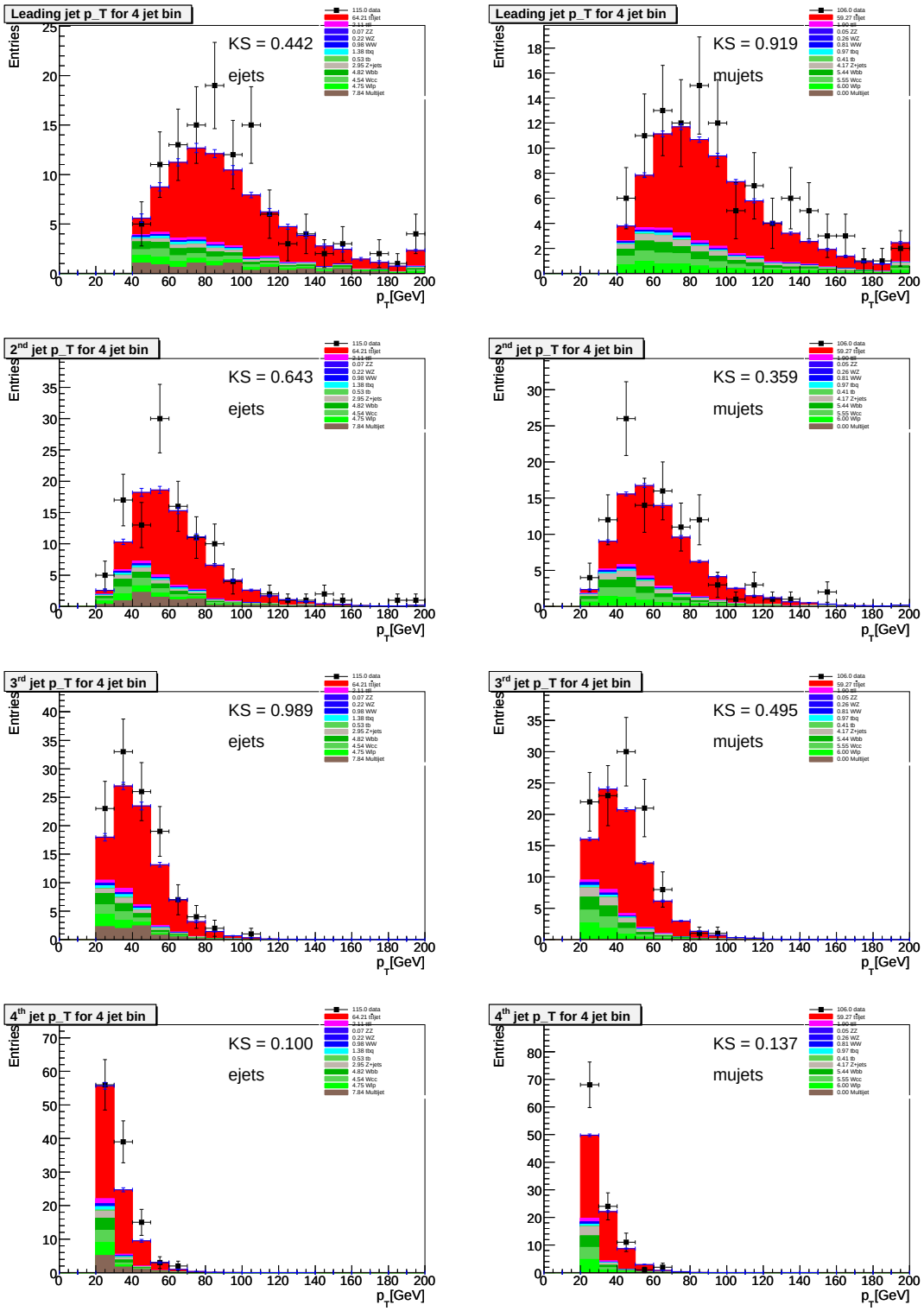


Figure C.14: Comparison for Run IIa between data and Monte Carlo for four or more jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

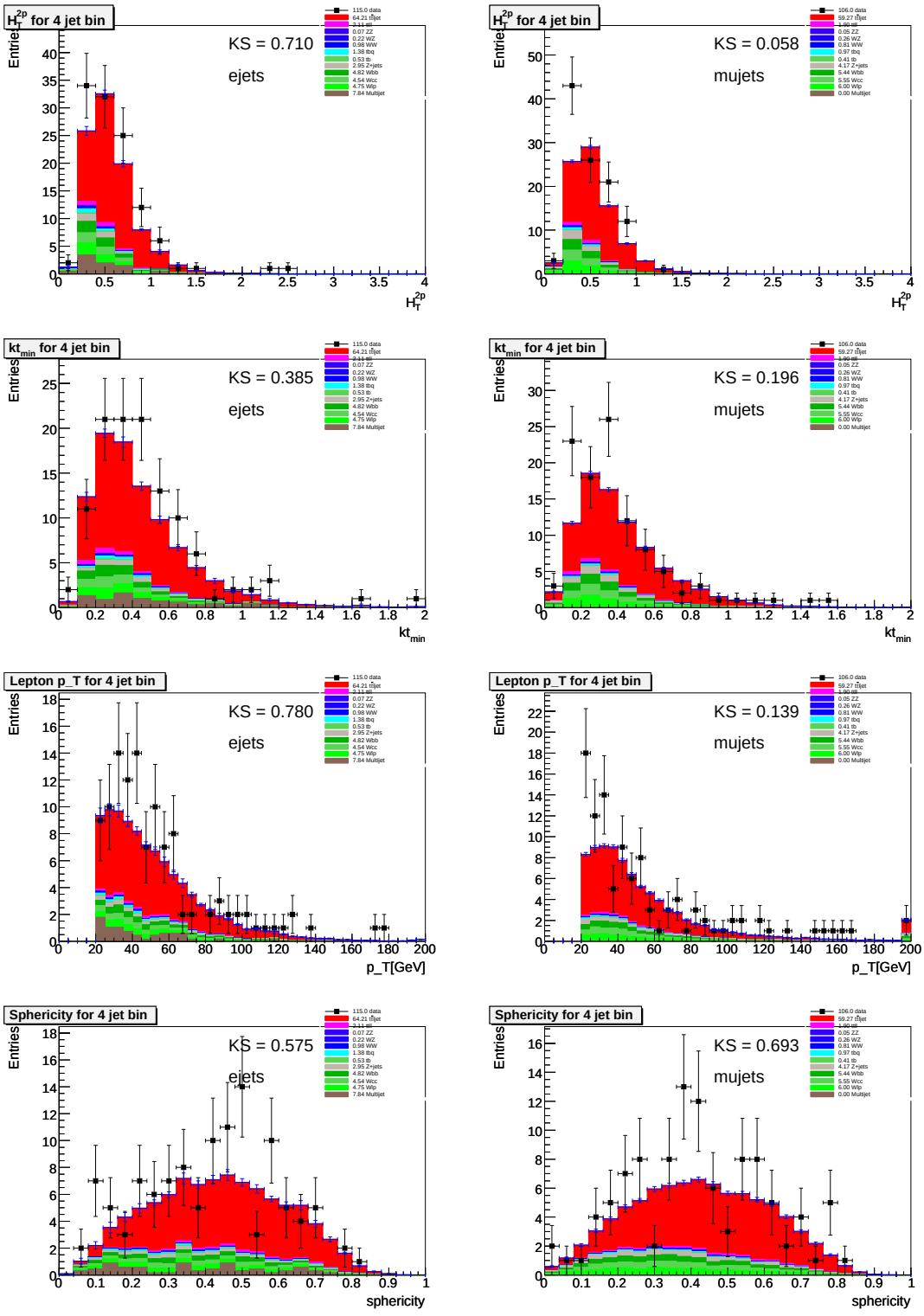


Figure C.15: Comparison for Run IIa between data and Monte Carlo for four or more jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

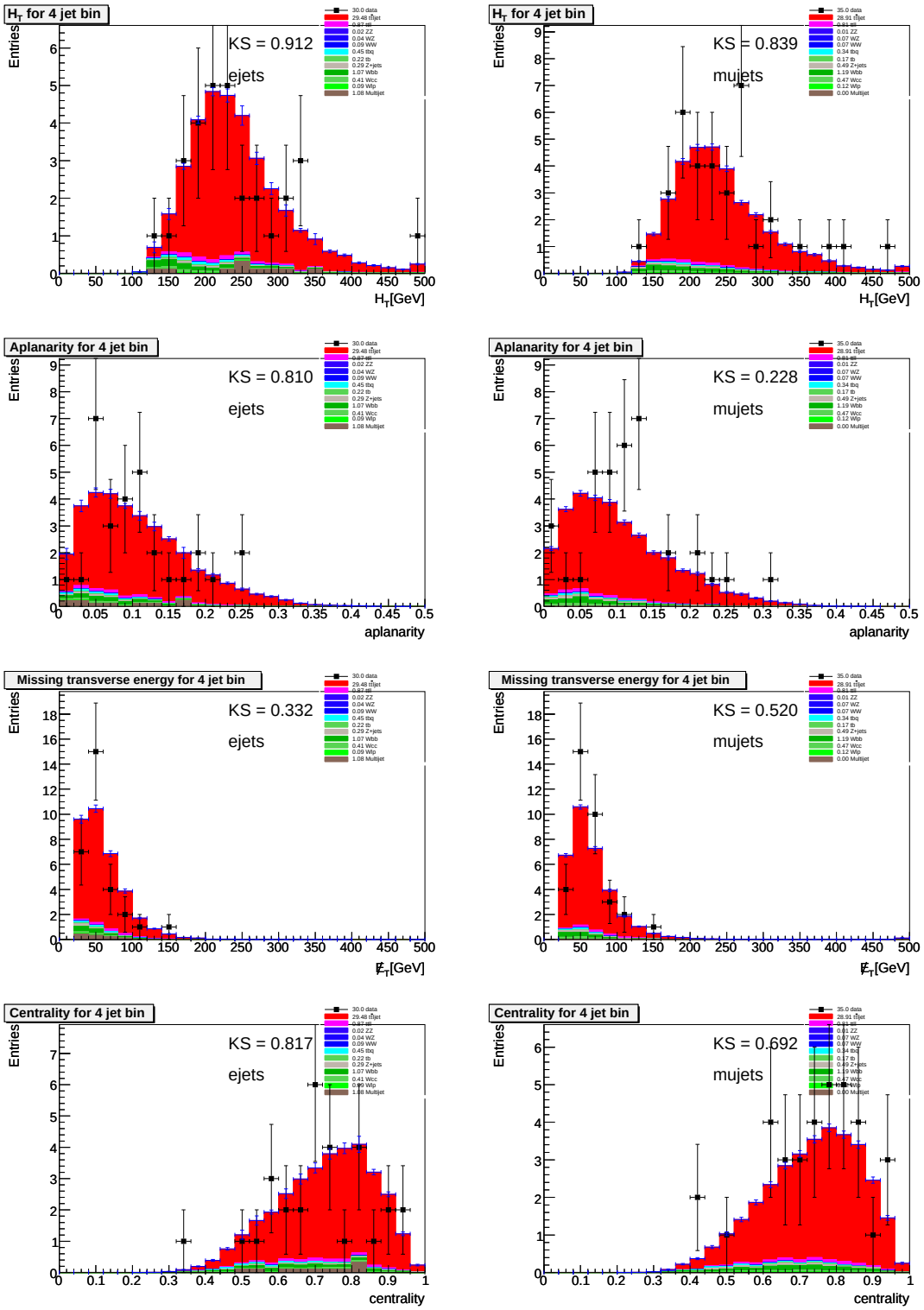


Figure C.16: Comparison for Run IIa between data and Monte Carlo for four or more jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

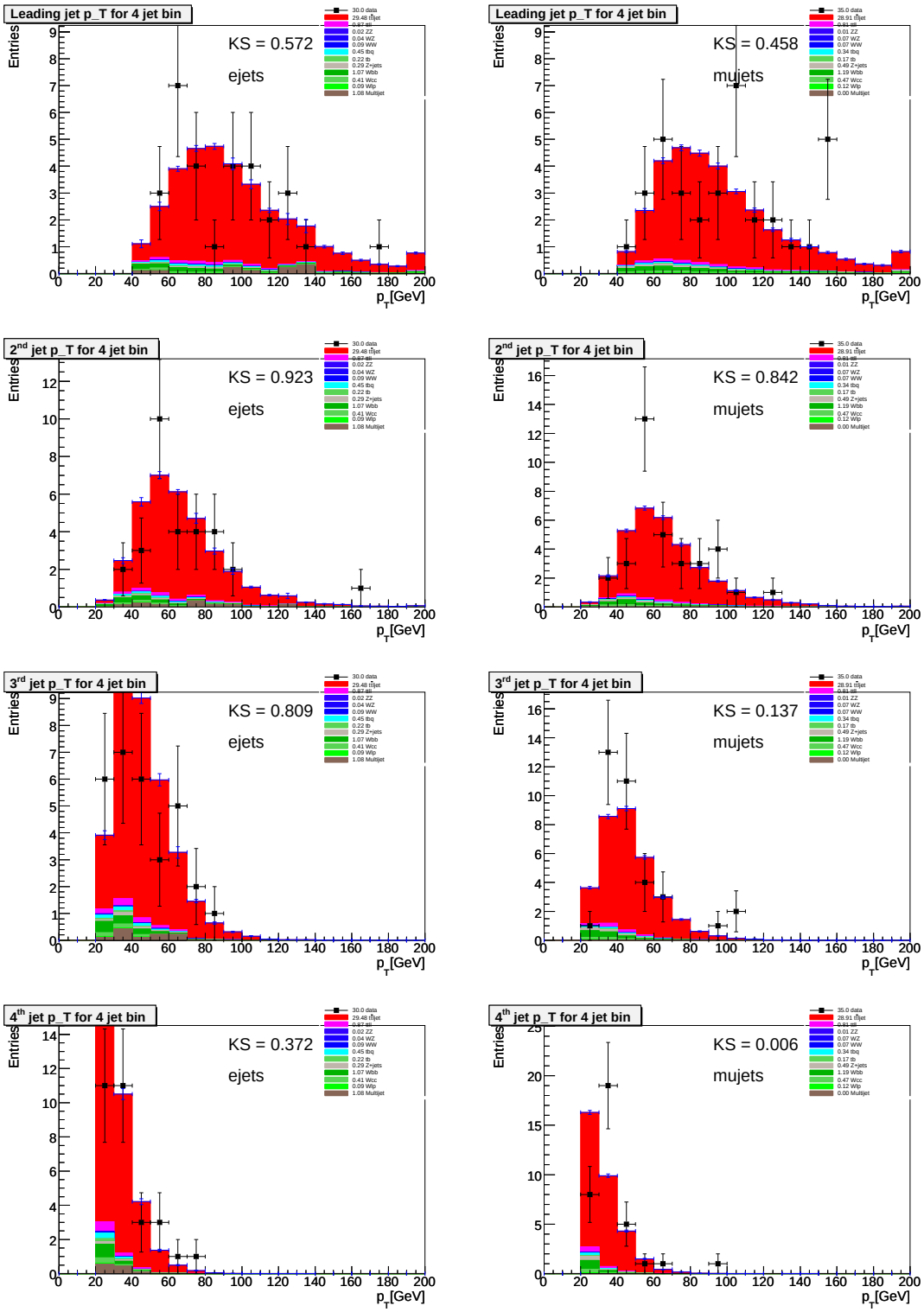


Figure C.17: Comparison for Run IIa between data and Monte Carlo for four or more jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

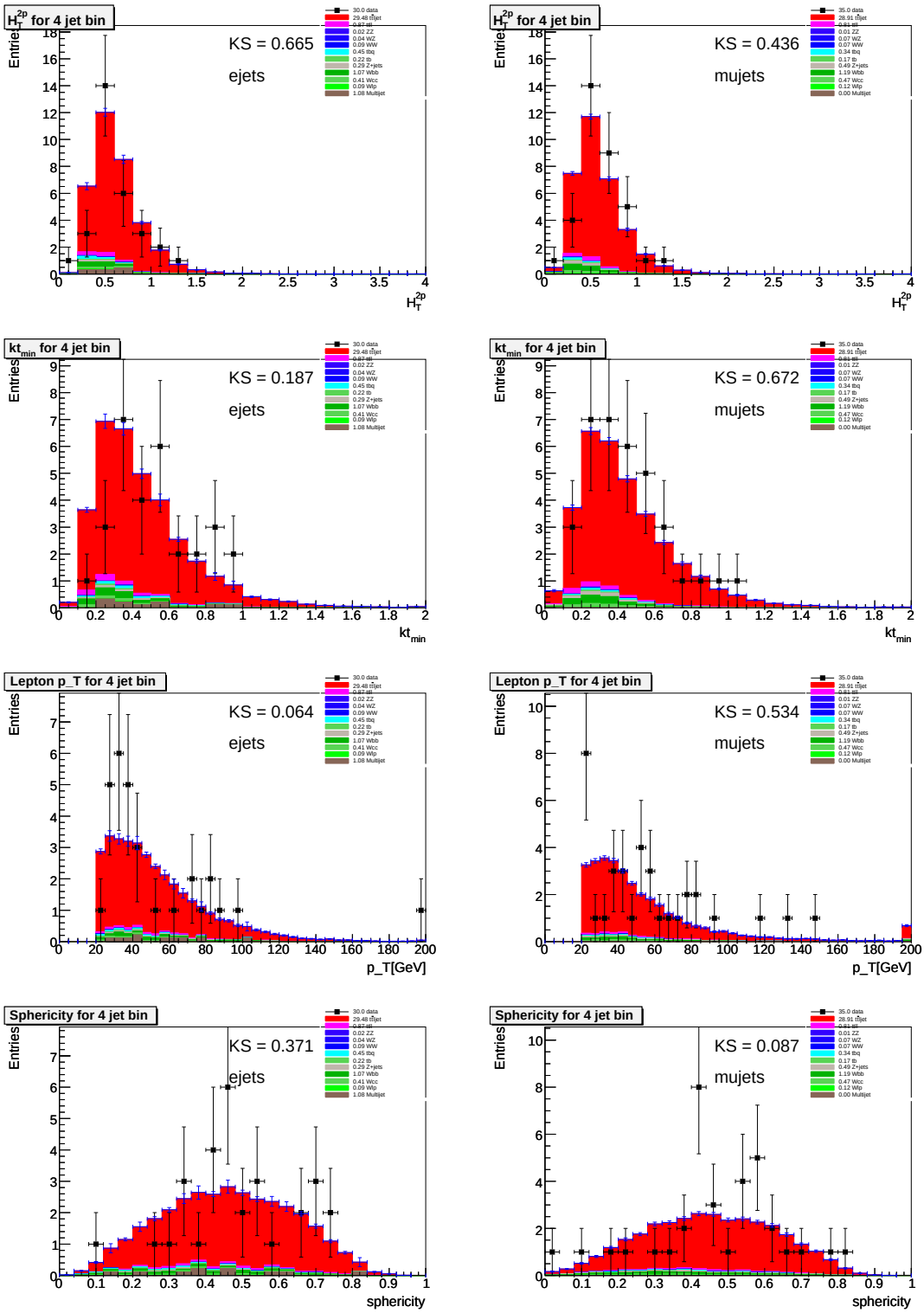


Figure C.18: Comparison for Run IIa between data and Monte Carlo for four or more jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

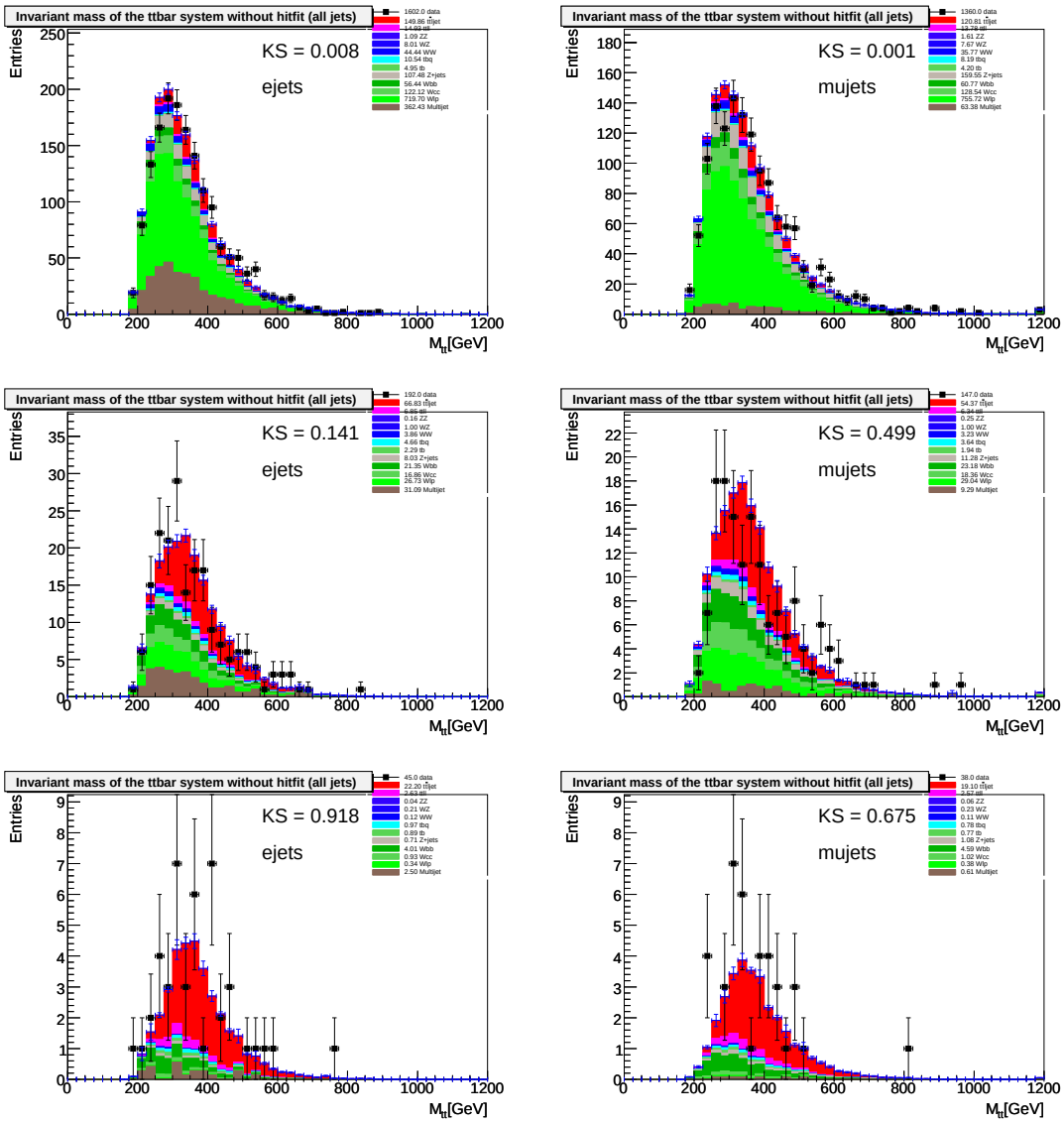


Figure C.19: Invariant mass distributions for three jet events in the e+jets (left), the μ+jets channel (right). First row corresponds to the selection before b-tagging. The second and third row show the results for $= 1$ and ≥ 2 medium NN b-tags, respectively.

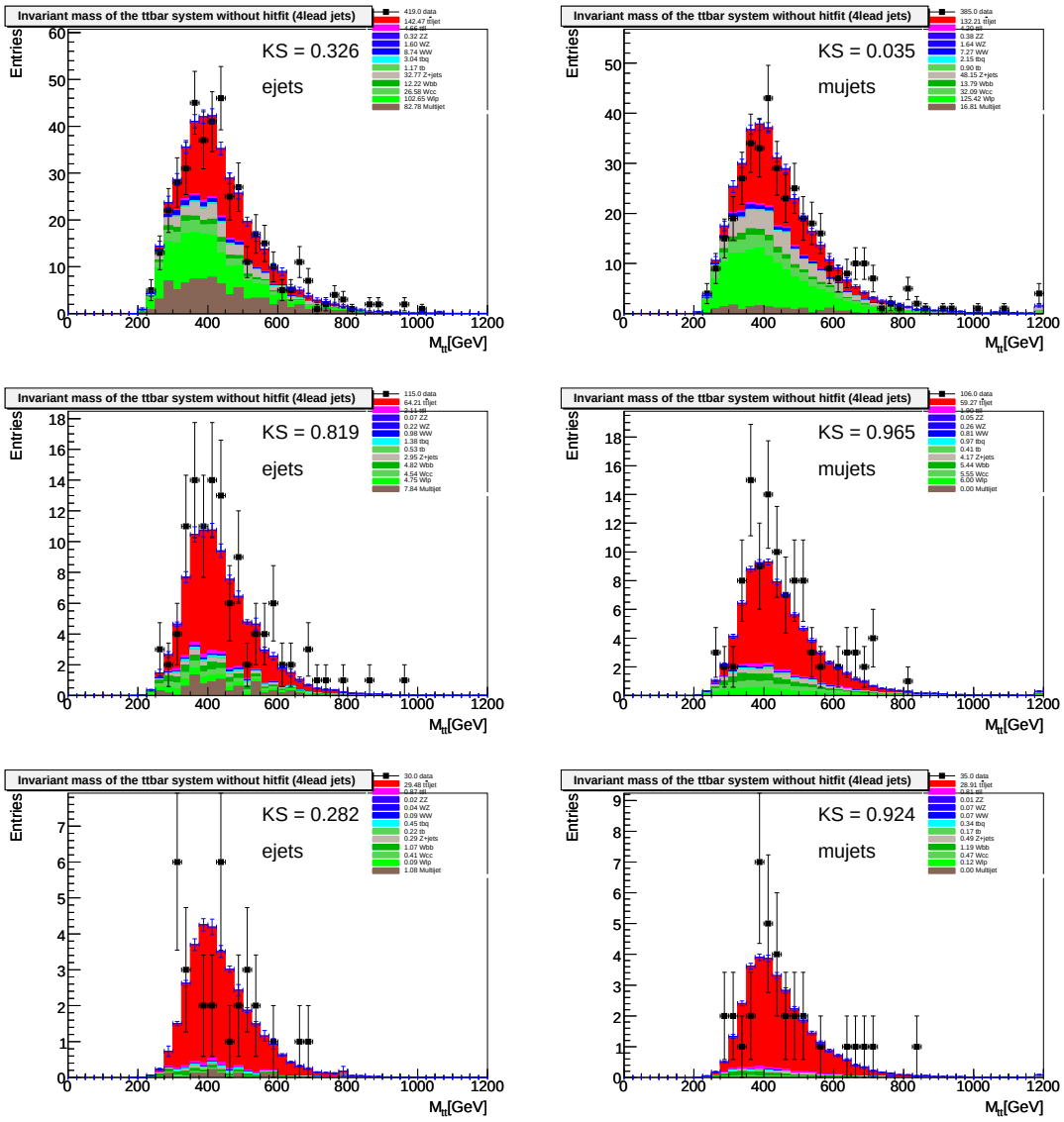


Figure C.20: Invariant mass distributions for four or more jet events in the e+jets (left), the μ+jets channel (right). First row corresponds to the selection before b-tagging. The second and third row show the results for $= 1$ and ≥ 2 medium NN b-tags, respectively.

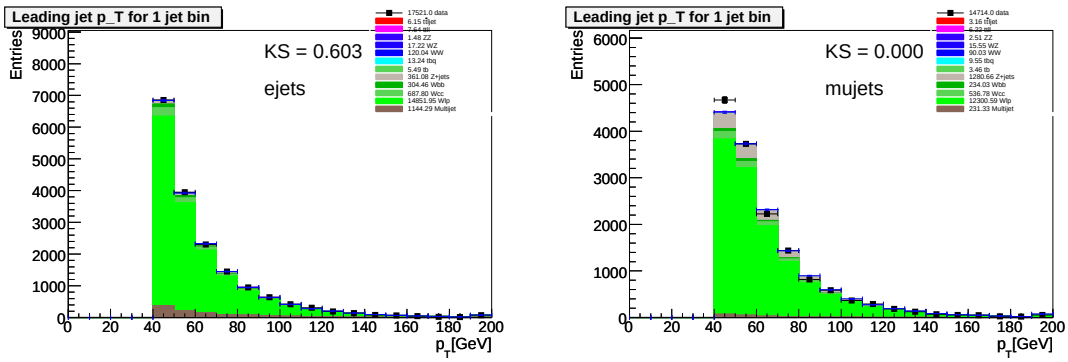


Figure C.21: Comparison for Run IIa between data and Monte Carlo with one jet before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables are: p_T of the leading jet.

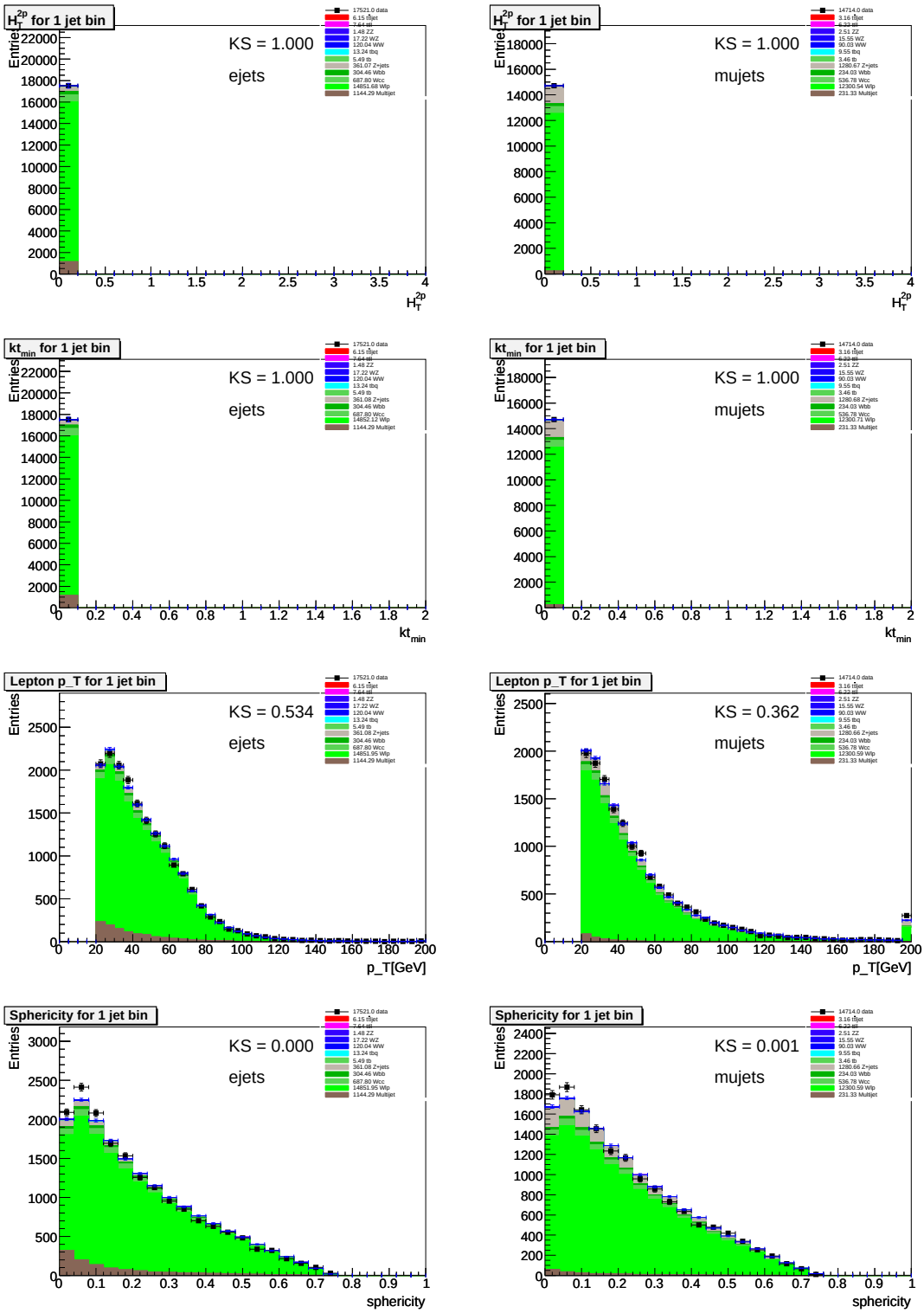


Figure C.22: Comparison for Run IIa between data and Monte Carlo with one jet before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

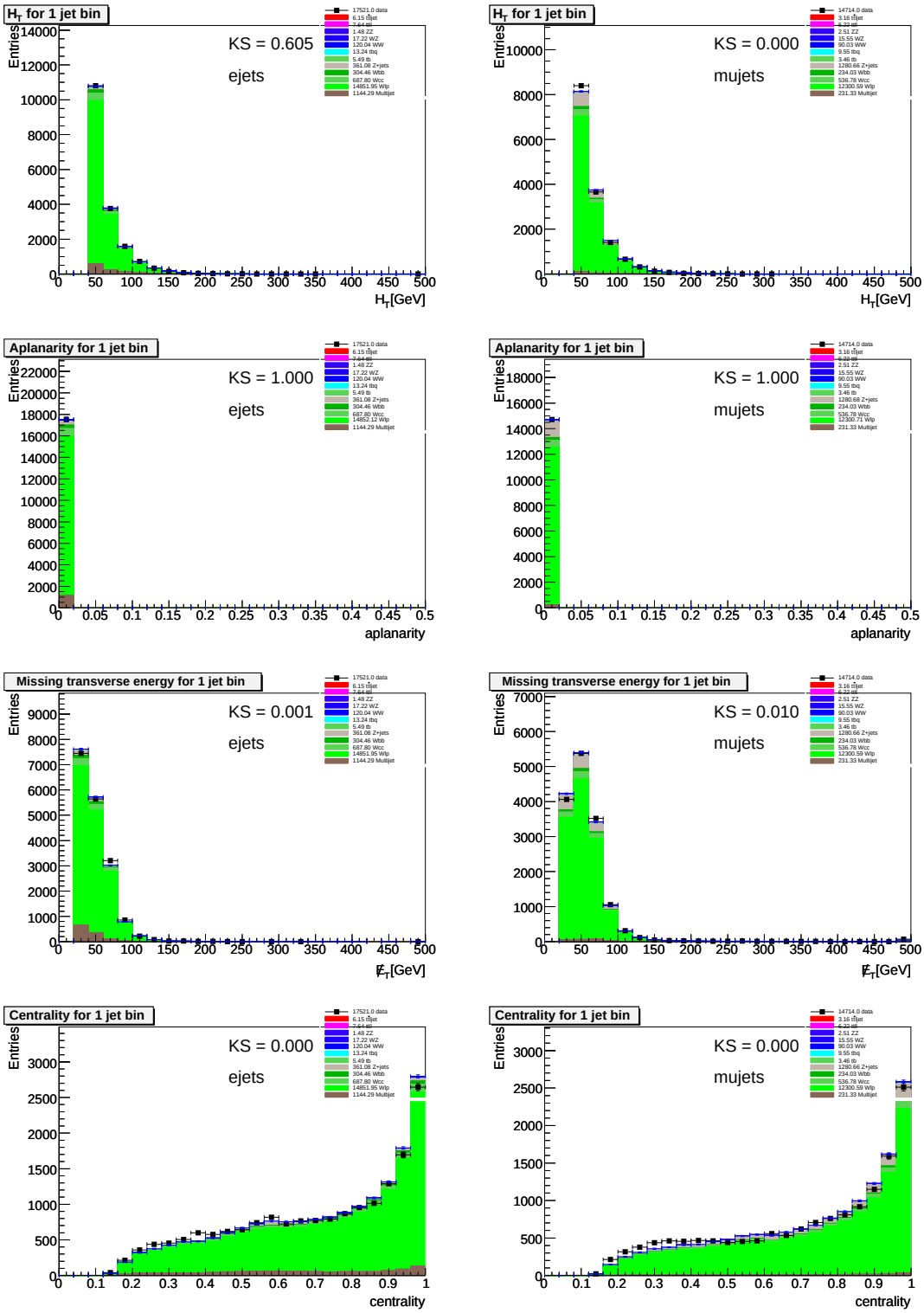


Figure C.23: Comparison for Run IIa between data and Monte Carlo with one jet before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

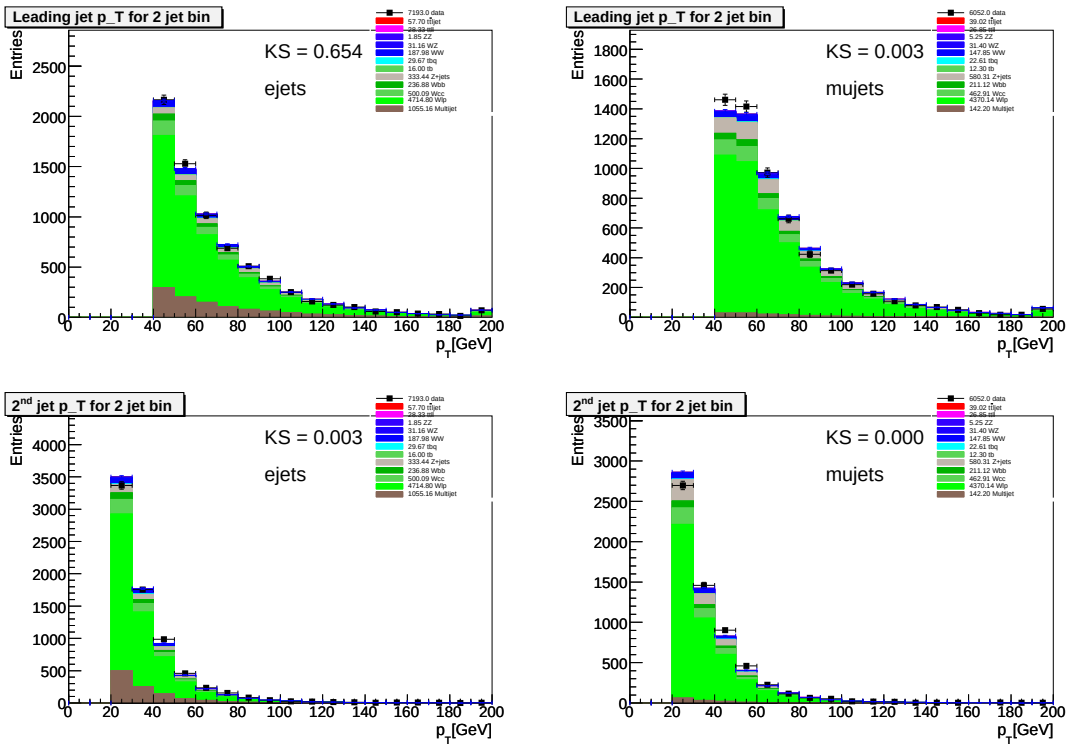


Figure C.24: Comparison for Run IIa between data and Monte Carlo with two jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet and p_T of the second leading jet.

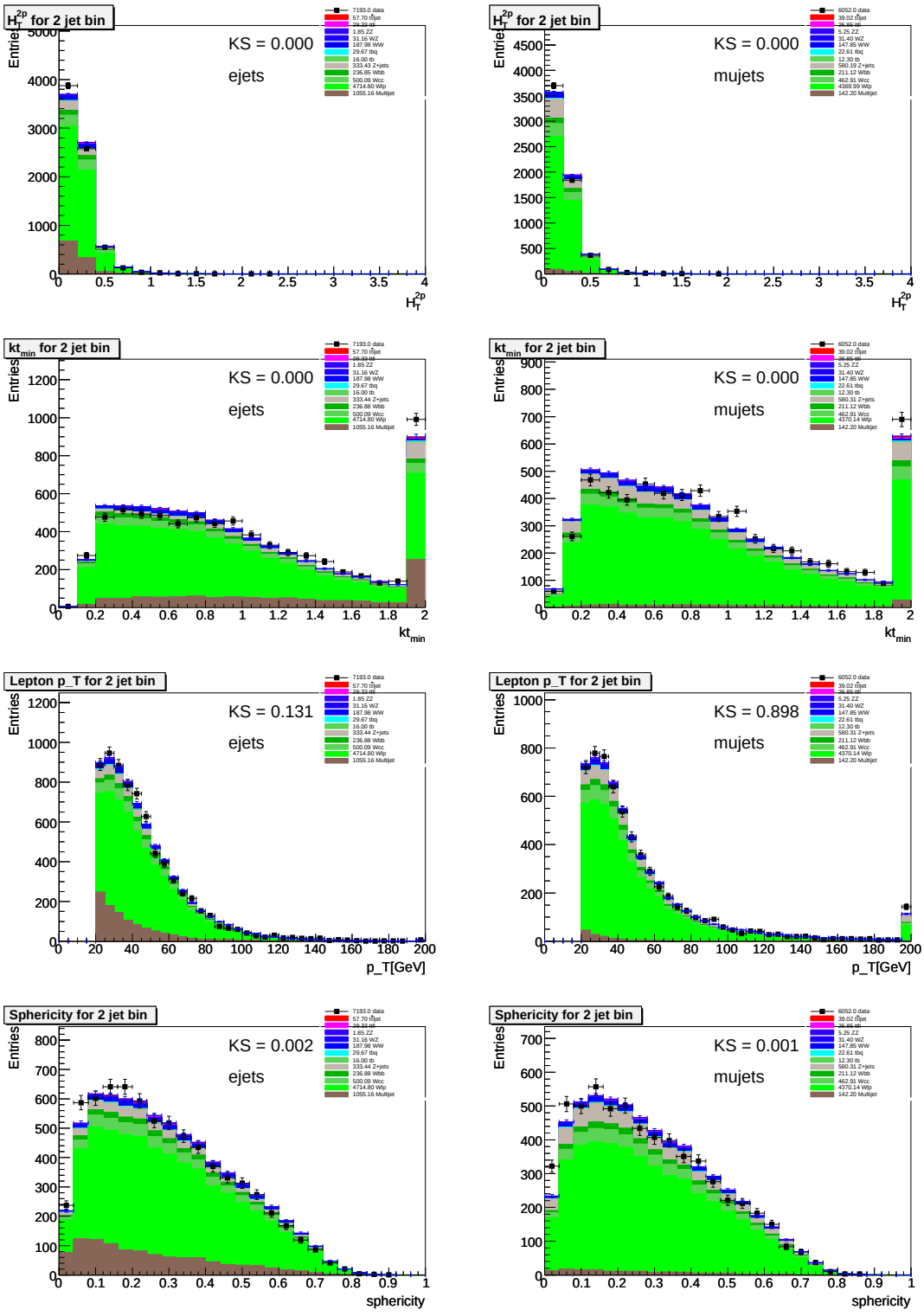


Figure C.25: Comparison for Run IIa between data and Monte Carlo with two jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

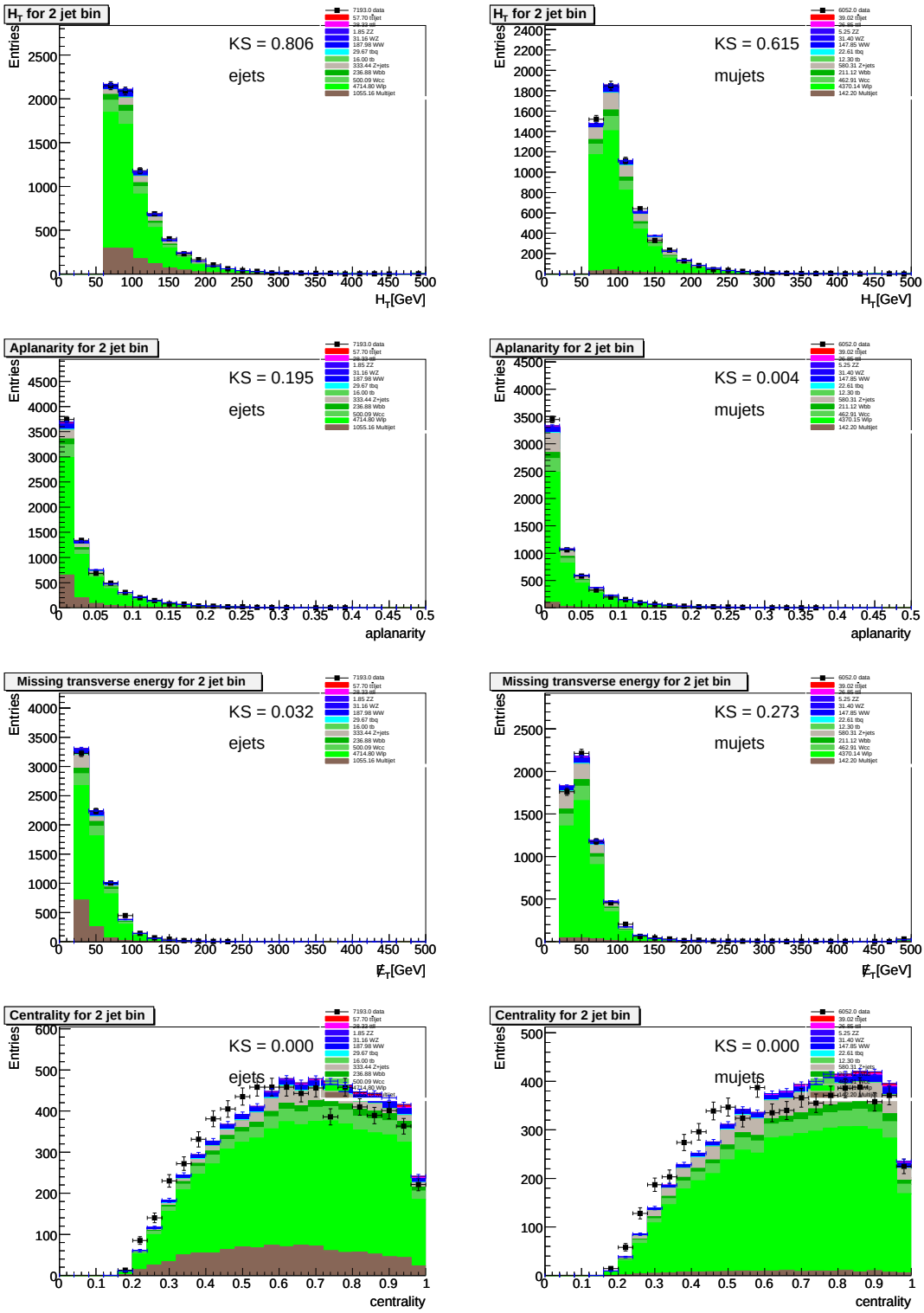


Figure C.26: Comparison for Run IIa between data and Monte Carlo with two jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

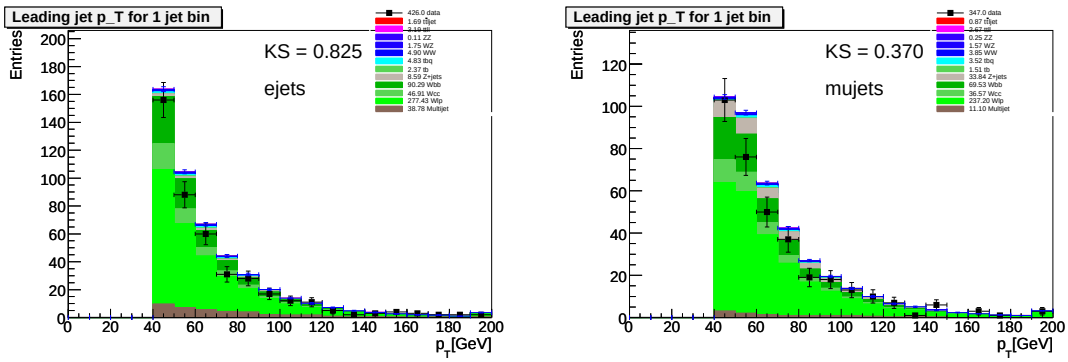


Figure C.27: Comparison for Run IIa between data and Monte Carlo with one jet including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables are: p_T of the leading jet.

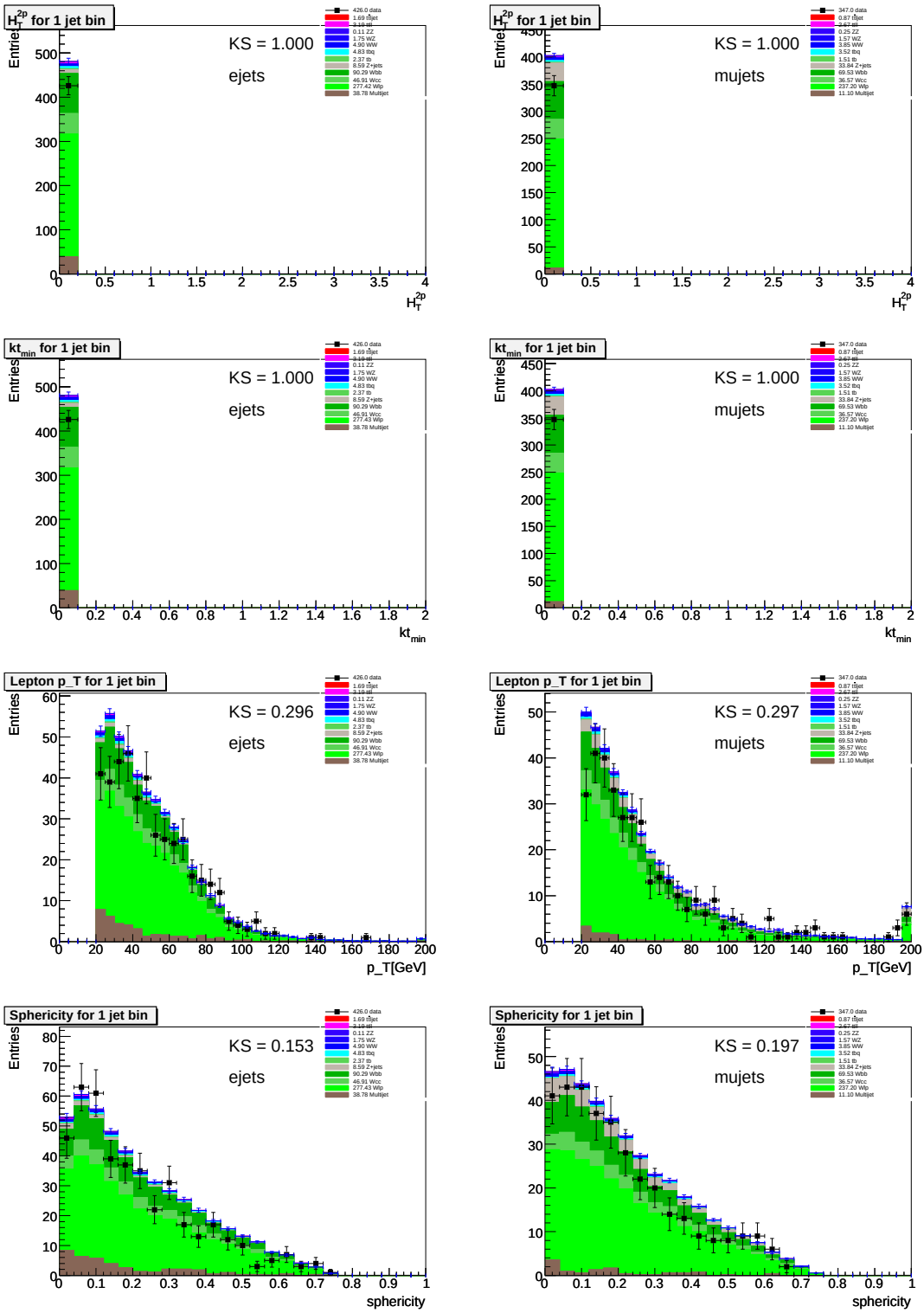


Figure C.28: Comparison for Run IIa between data and Monte Carlo with one jet including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

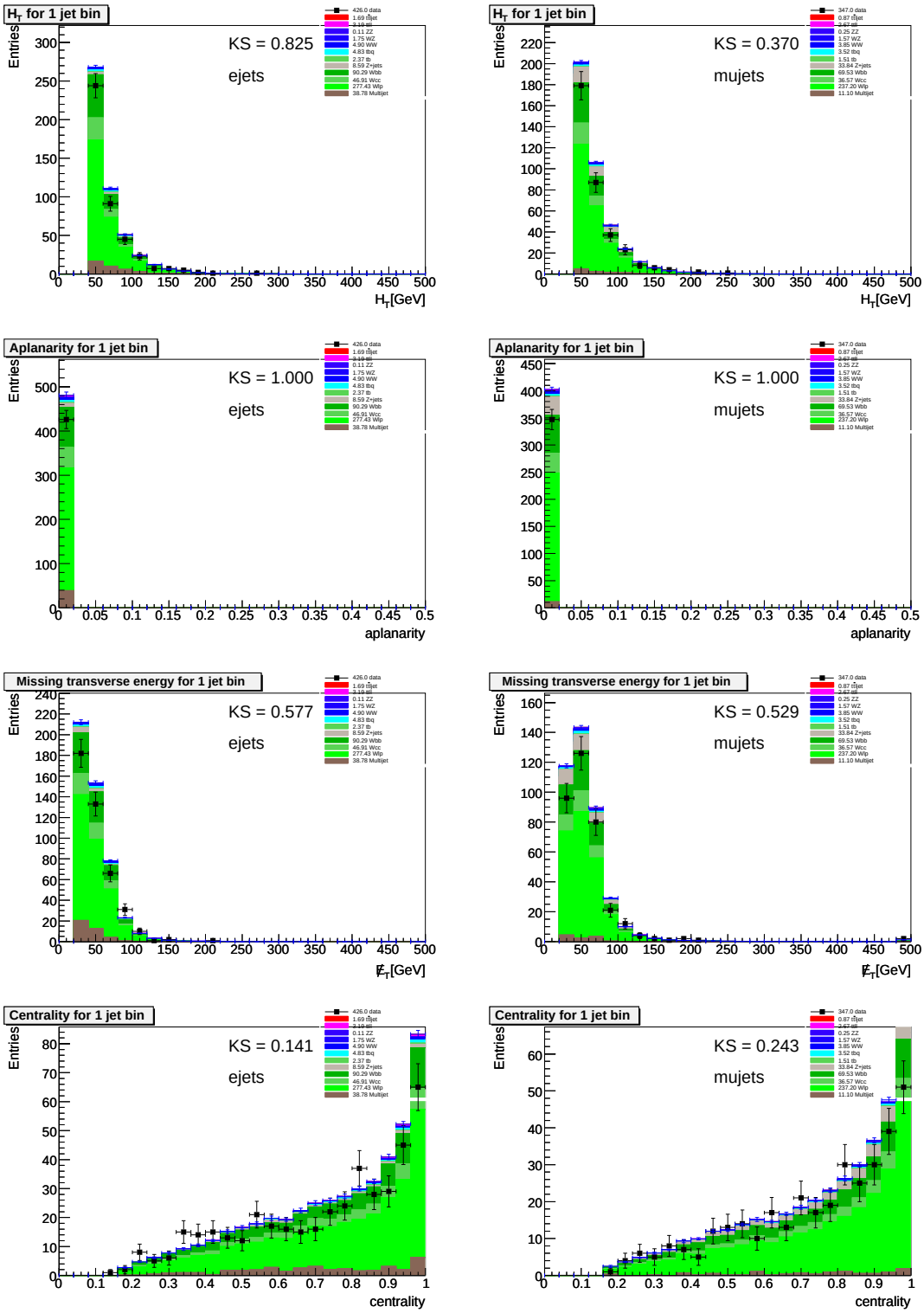


Figure C.29: Comparison for Run IIa between data and Monte Carlo with one jet including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

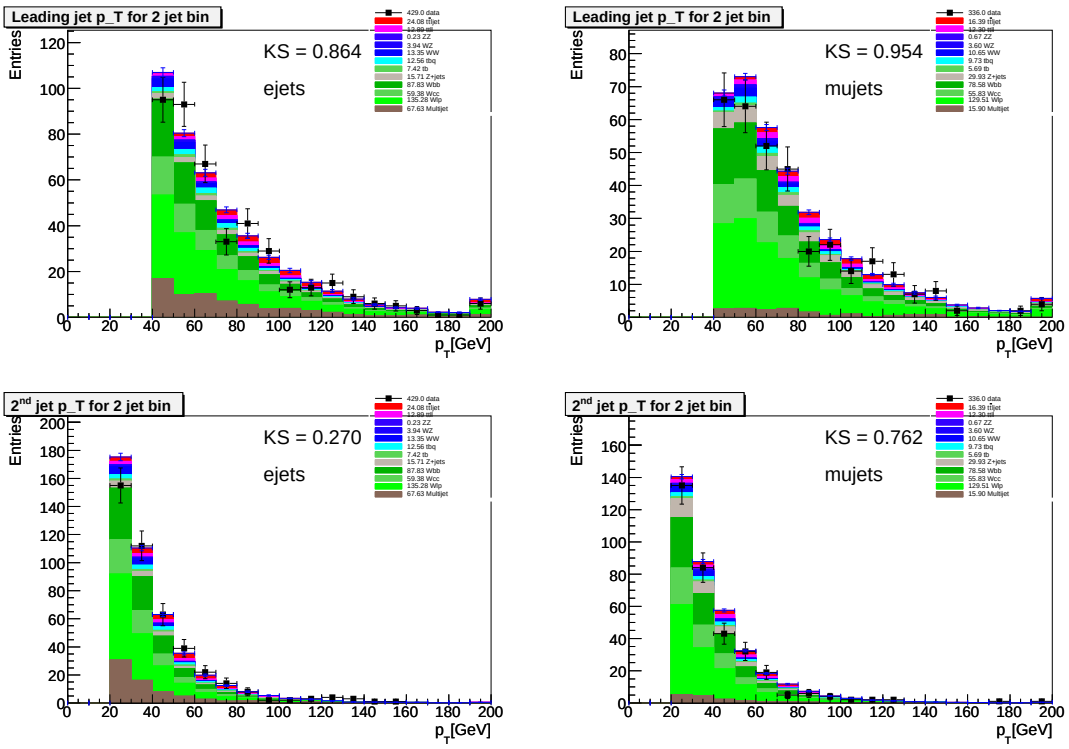


Figure C.30: Comparison for Run IIa between data and Monte Carlo with two jets including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet and p_T of the second leading jet.

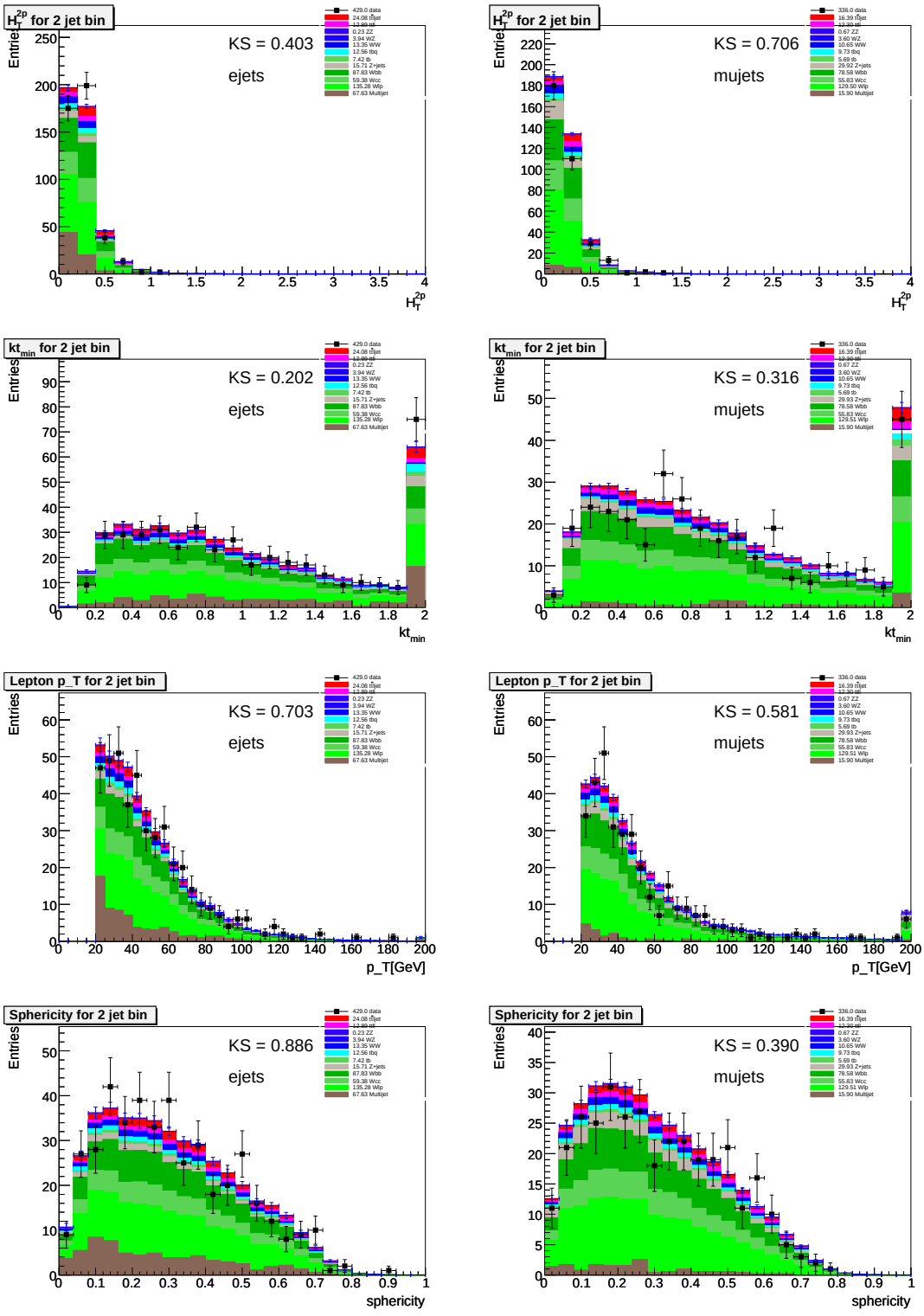


Figure C.31: Comparison for Run IIa between data and Monte Carlo with two jets including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

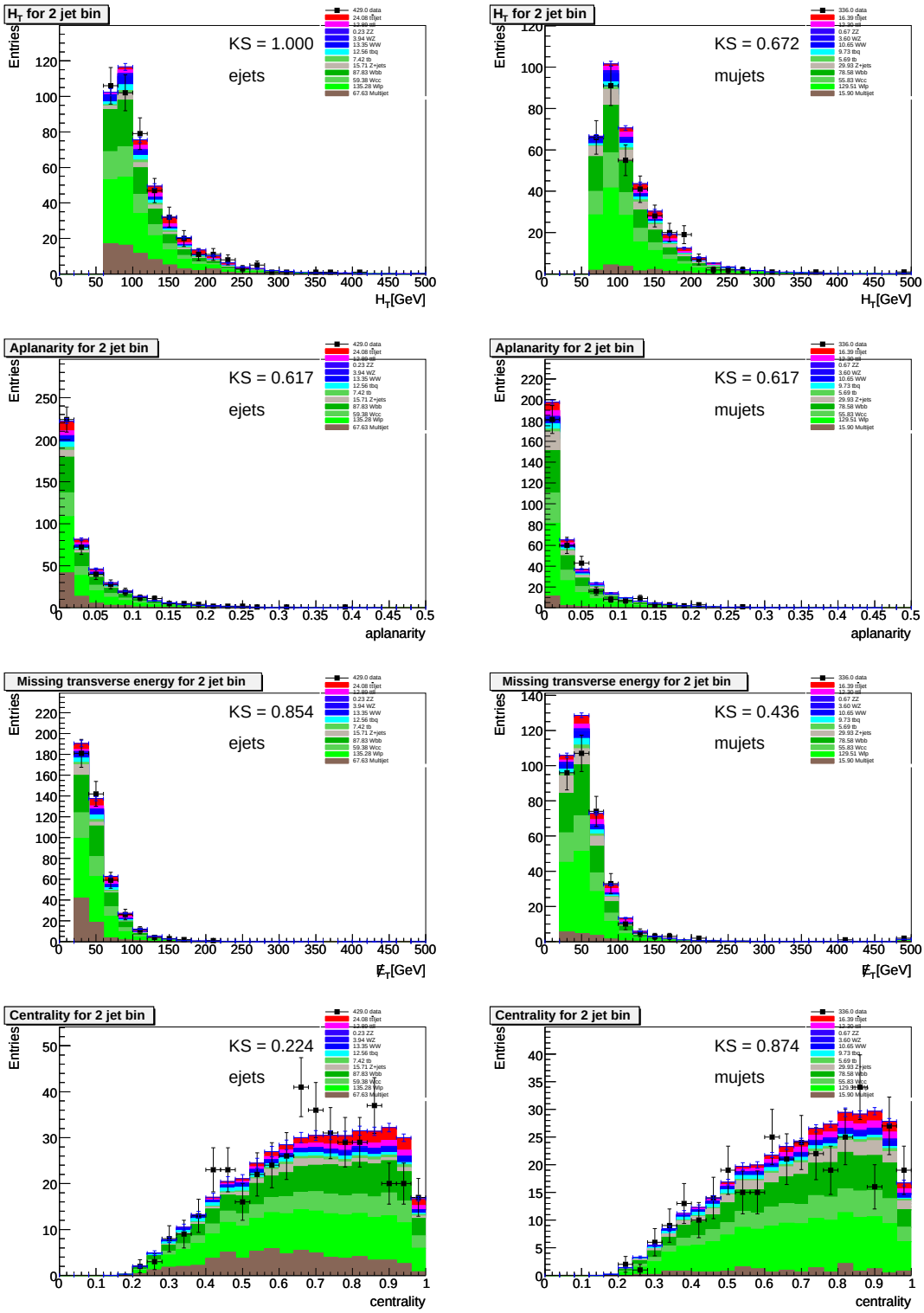


Figure C.32: Comparison for Run IIa between data and Monte Carlo with two jets including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

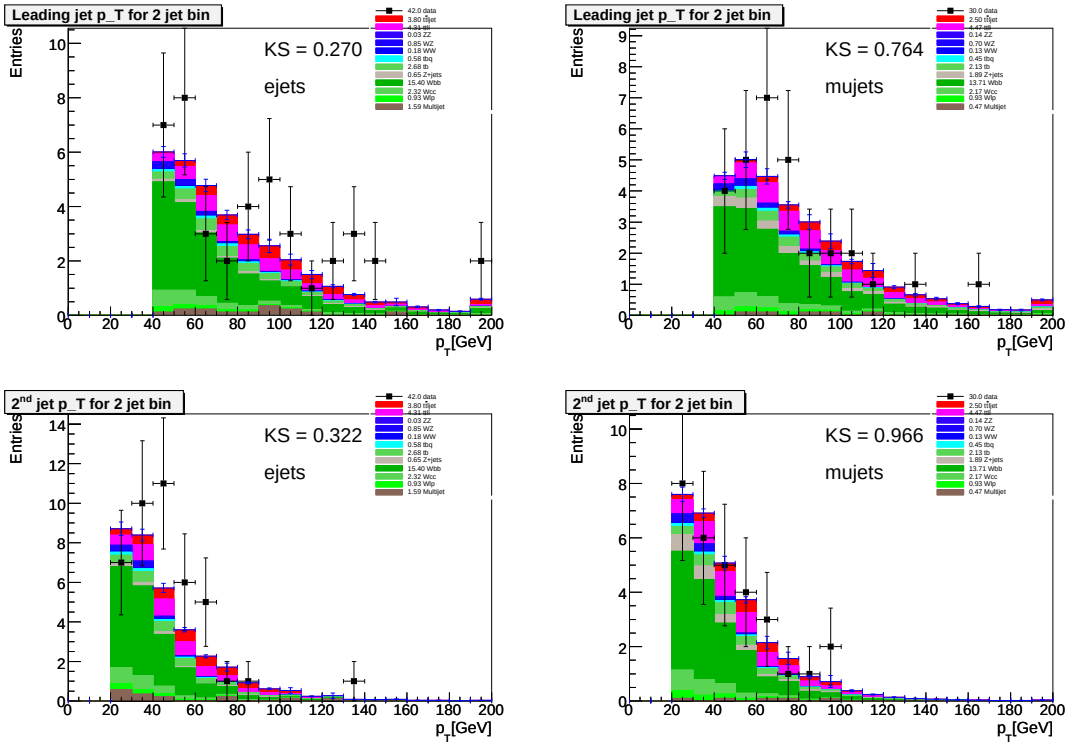


Figure C.33: Comparison for Run IIa between data and Monte Carlo with two jets including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet and p_T of the second leading jet.

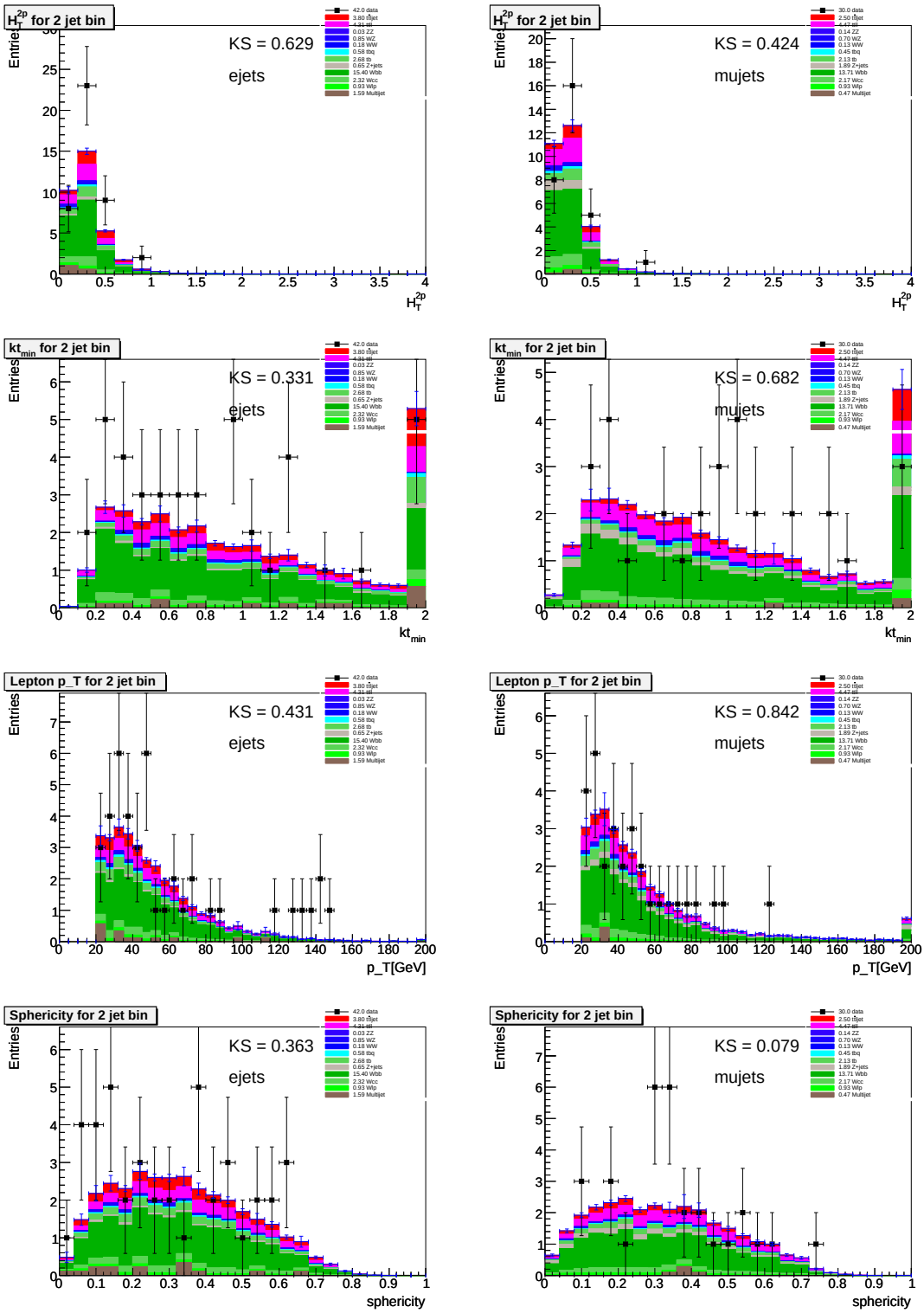


Figure C.34: Comparison for Run IIa between data and Monte Carlo with two jets including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

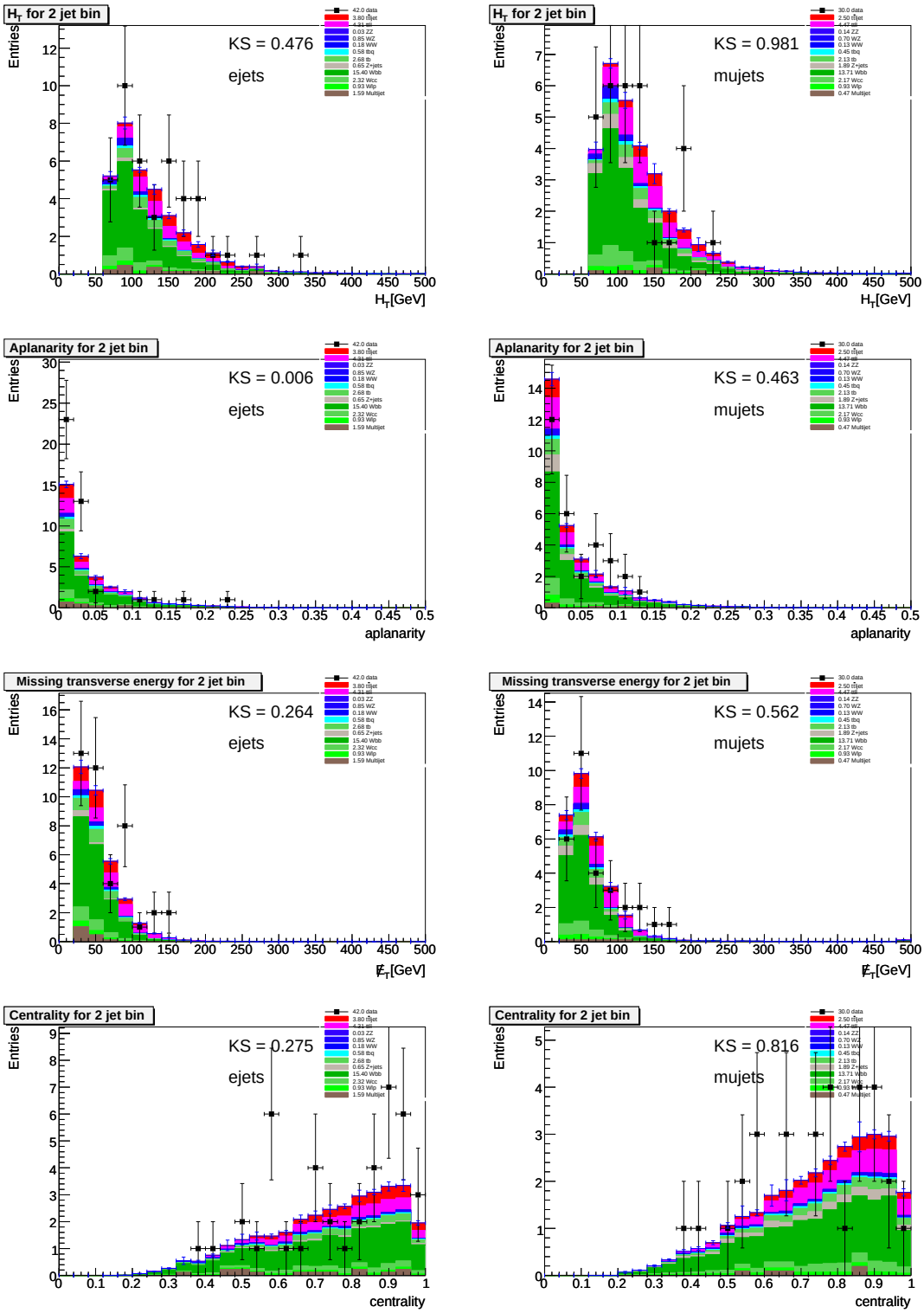


Figure C.35: Comparison for Run IIa between data and Monte Carlo with two jets including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

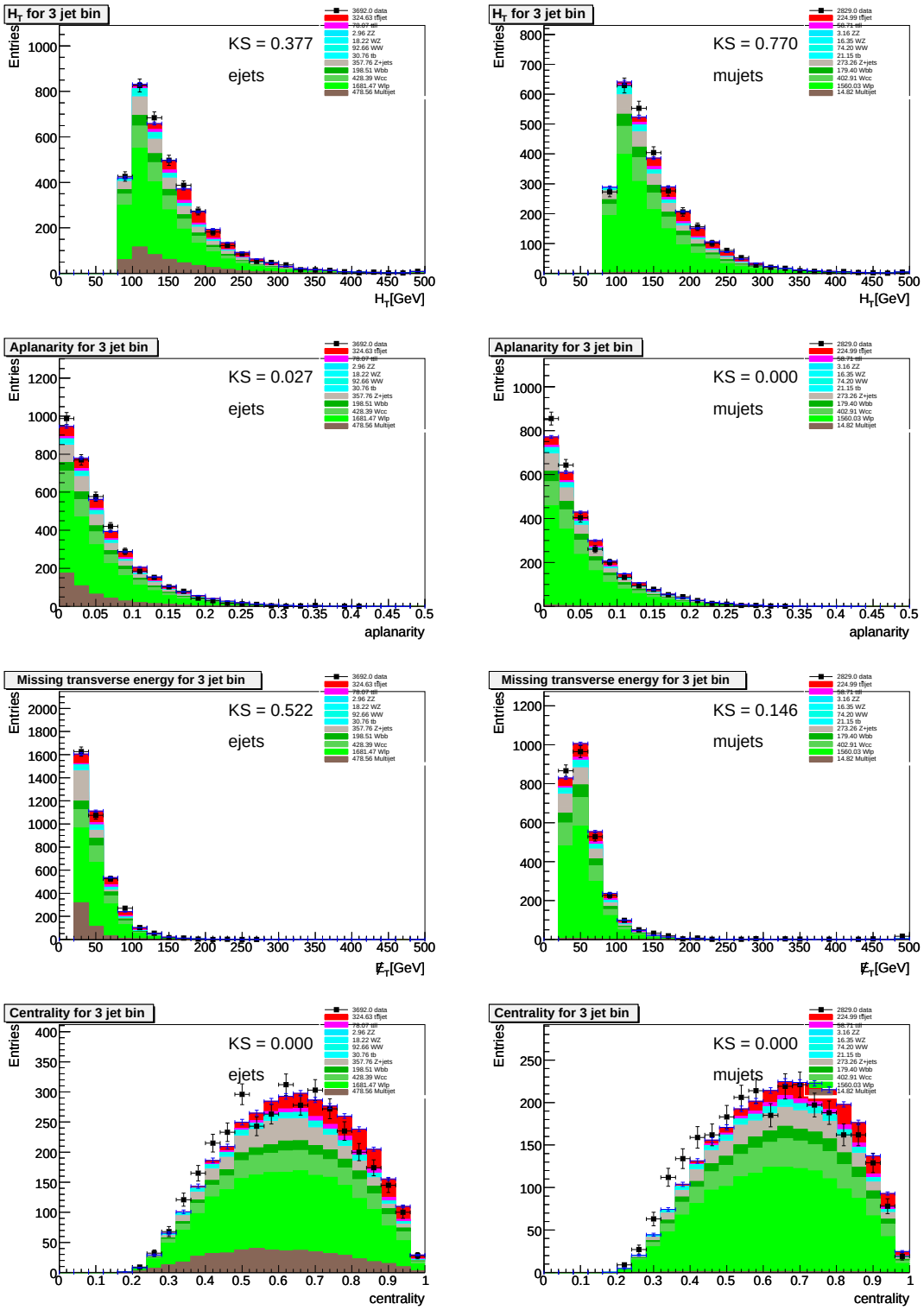


Figure C.36: Comparison between data and Monte Carlo with three jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

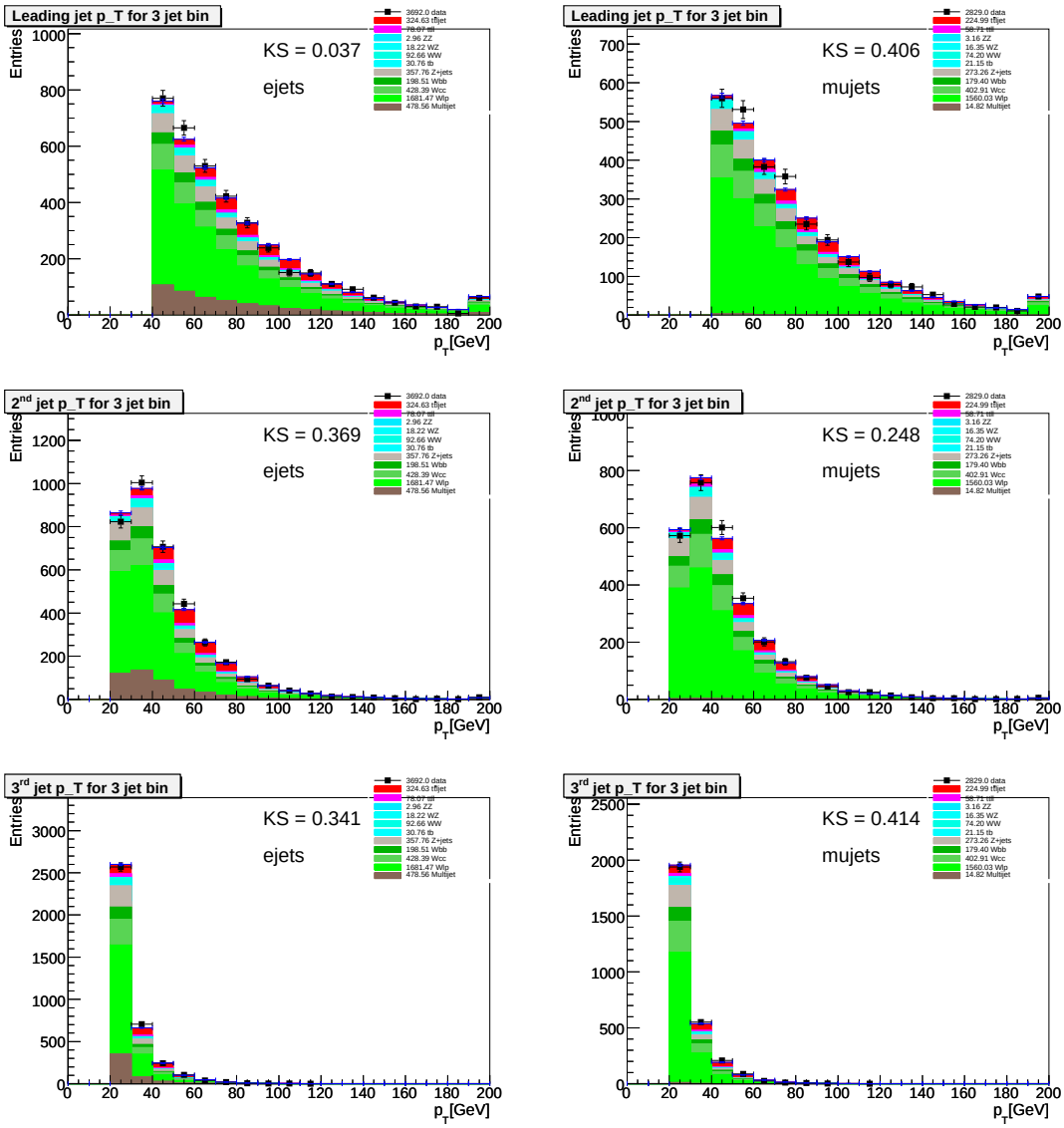


Figure C.37: Comparison for Run IIb between data and Monte Carlo with three jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

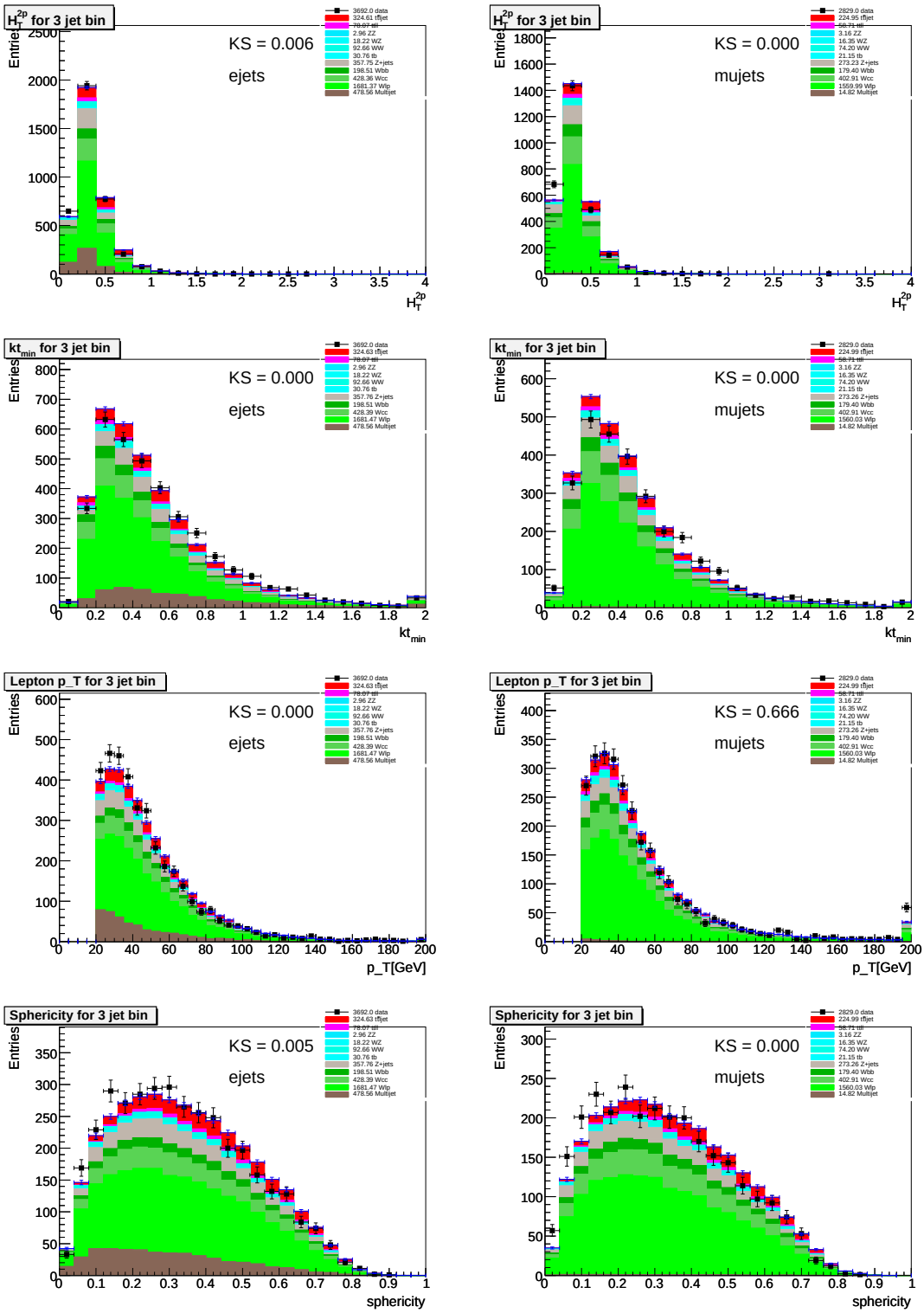


Figure C.38: Comparison for Run IIb between data and Monte Carlo with three jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Kt-minp, the p_T of the lepton, and sphericity.

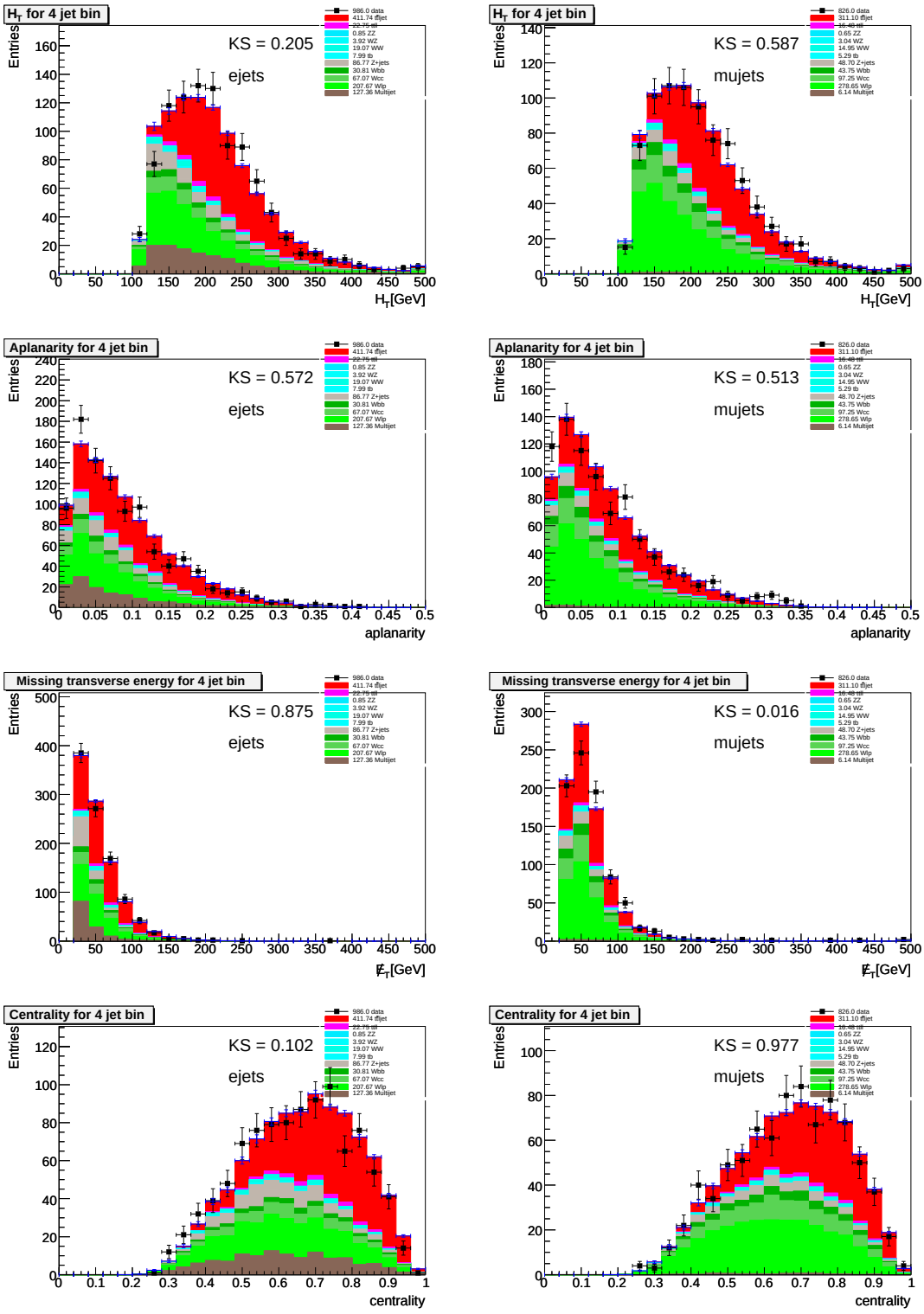


Figure C.39: Comparison for Run IIb between data and Monte Carlo with four or more jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

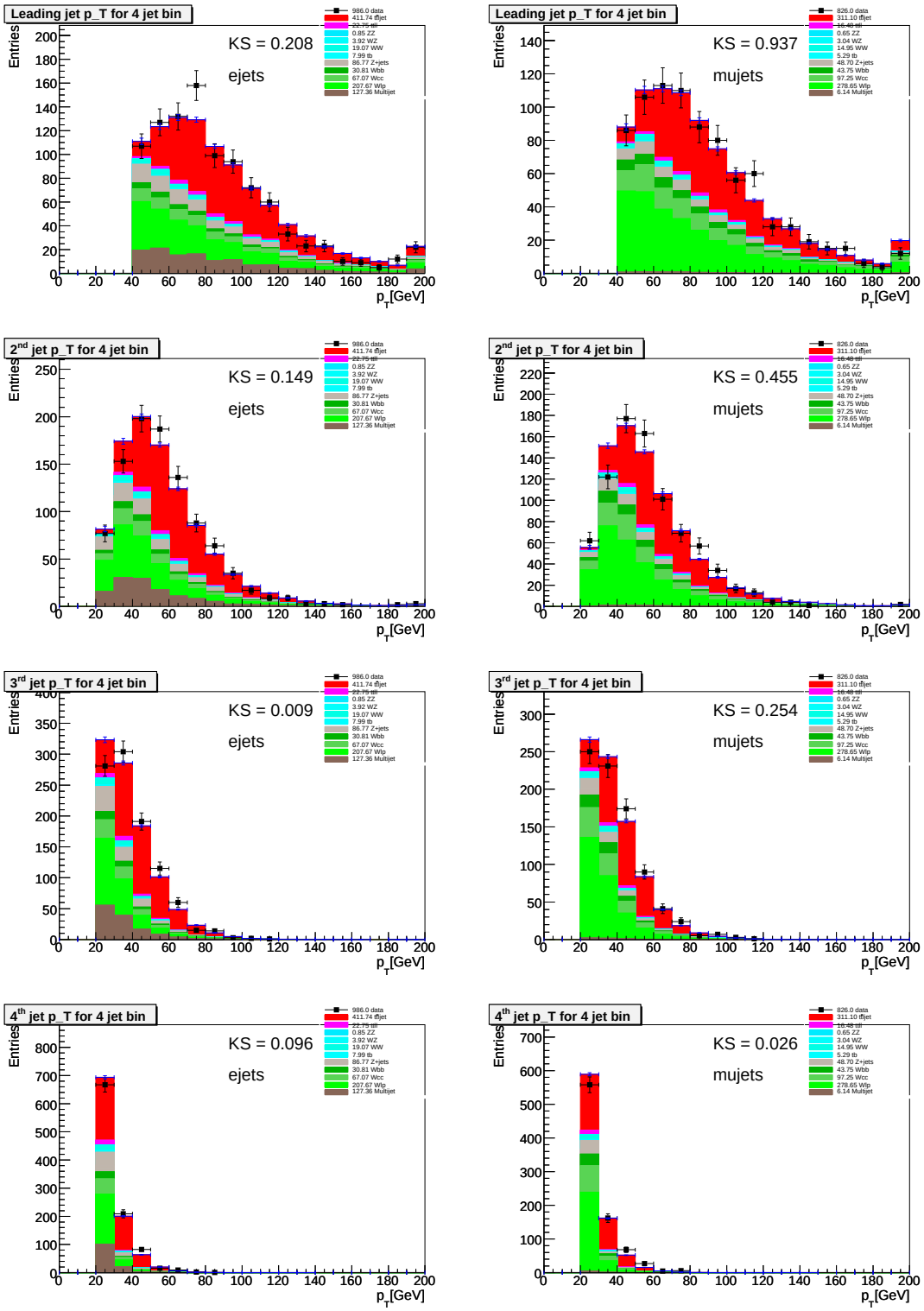


Figure C.40: Comparison for Run IIb between data and Monte Carlo with four or more jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

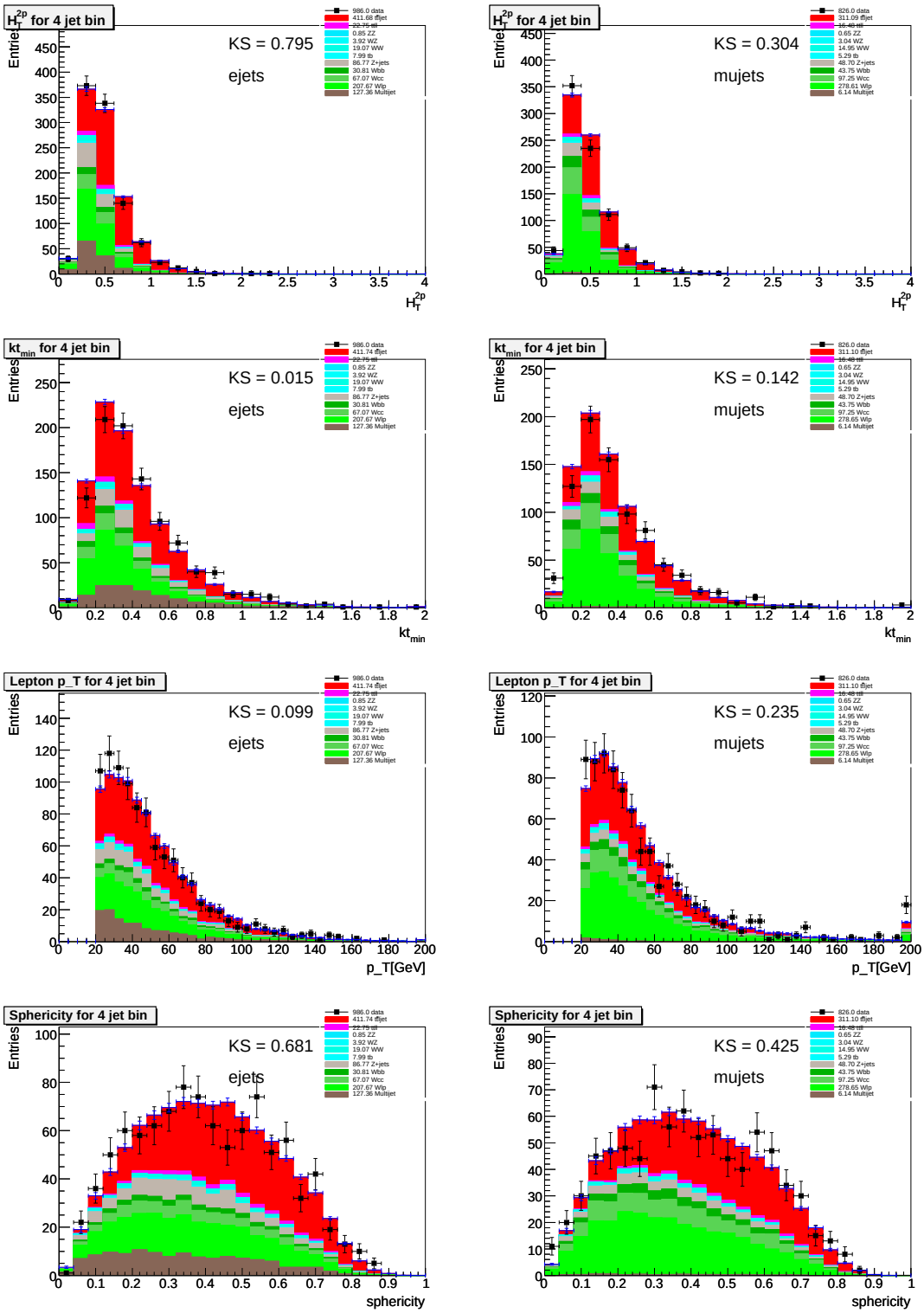


Figure C.41: Comparison for Run IIb between data and Monte Carlo with four or more jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

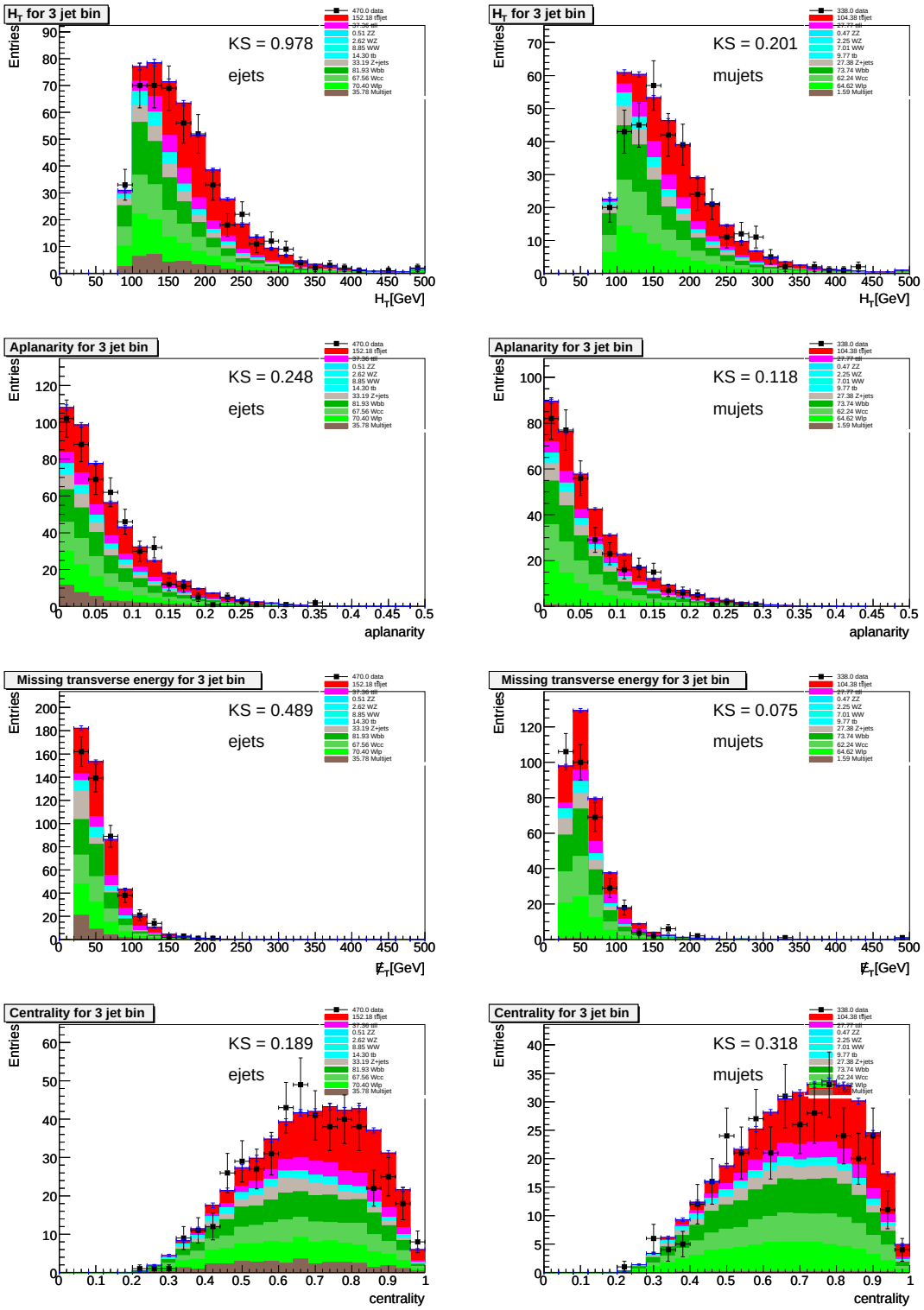


Figure C.42: Comparison for Run IIb between data and Monte Carlo for three jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

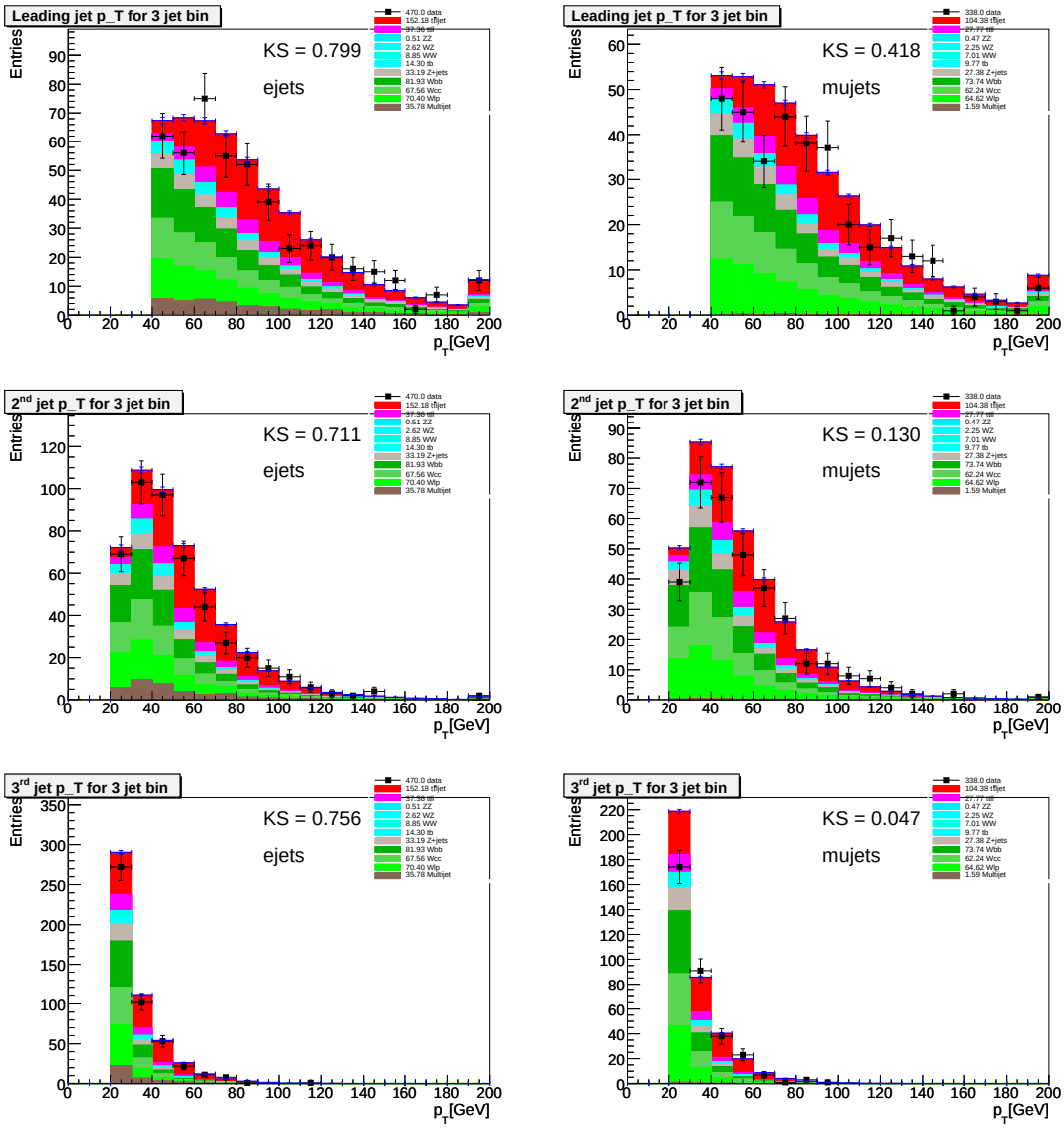


Figure C.43: Comparison for Run IIb between data and Monte Carlo for three jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet and p_T of the third leading jet.

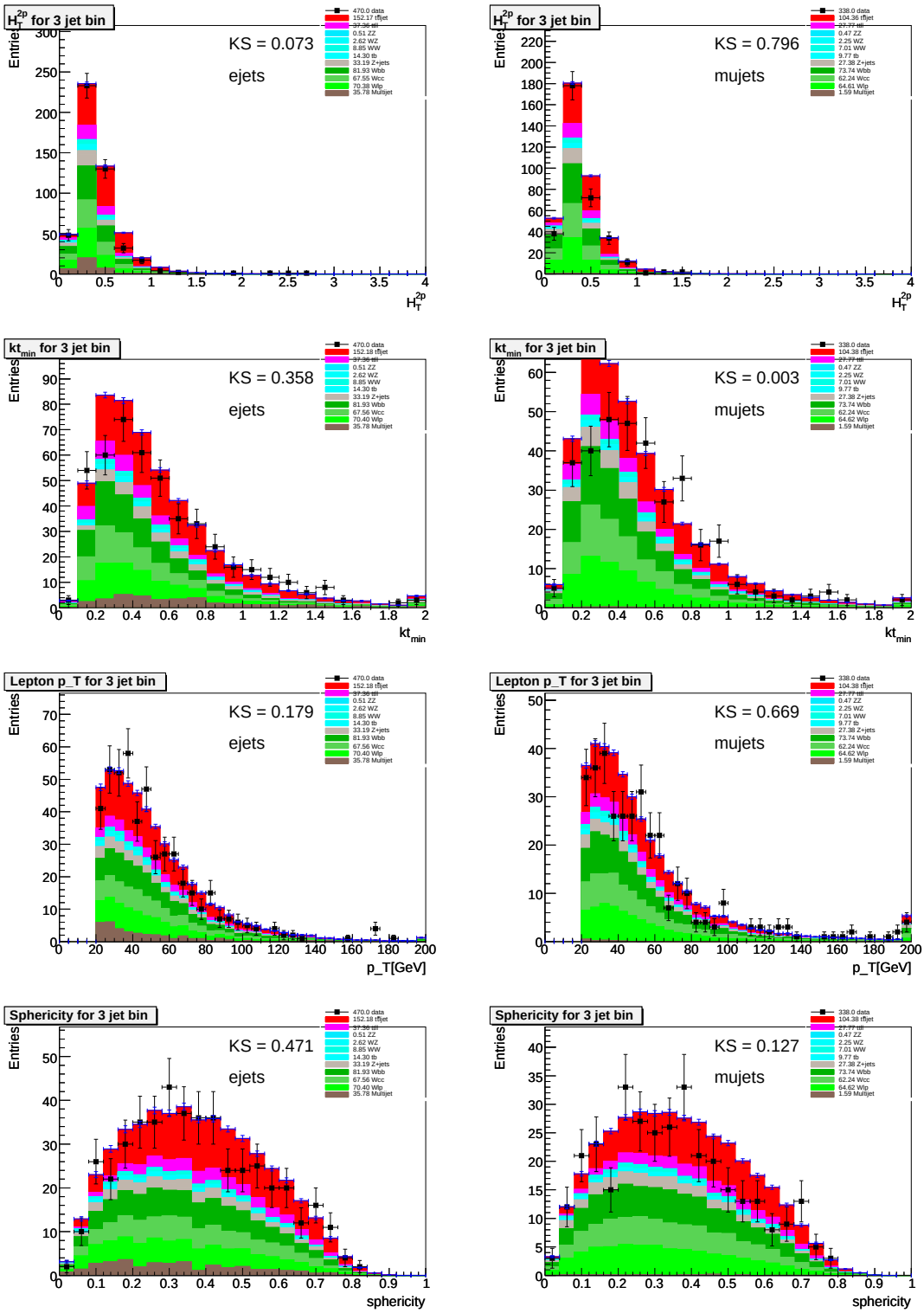


Figure C.44: Comparison for Run IIb between data and Monte Carlo for three jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

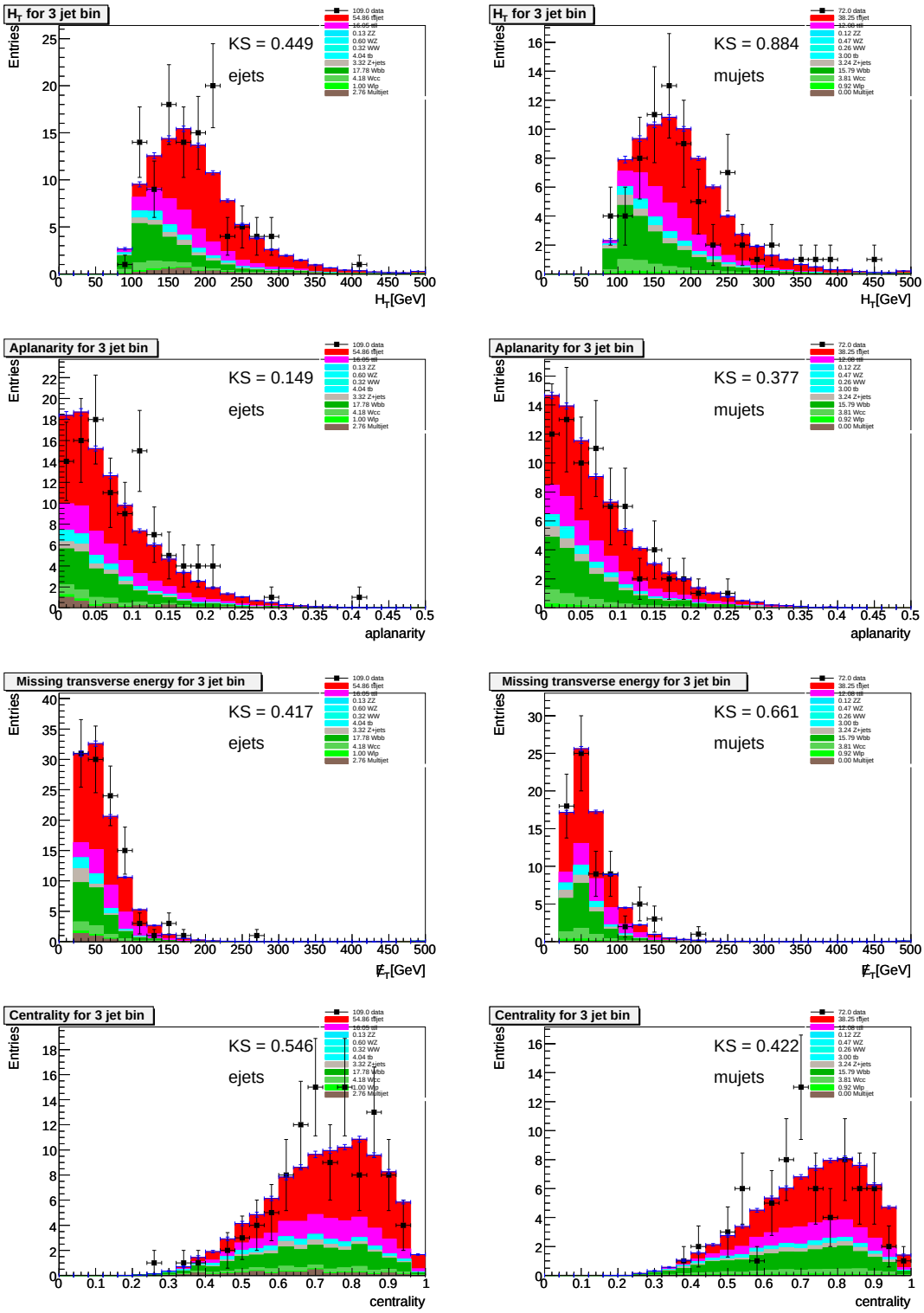


Figure C.45: Comparison for Run IIb between data and Monte Carlo for three jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

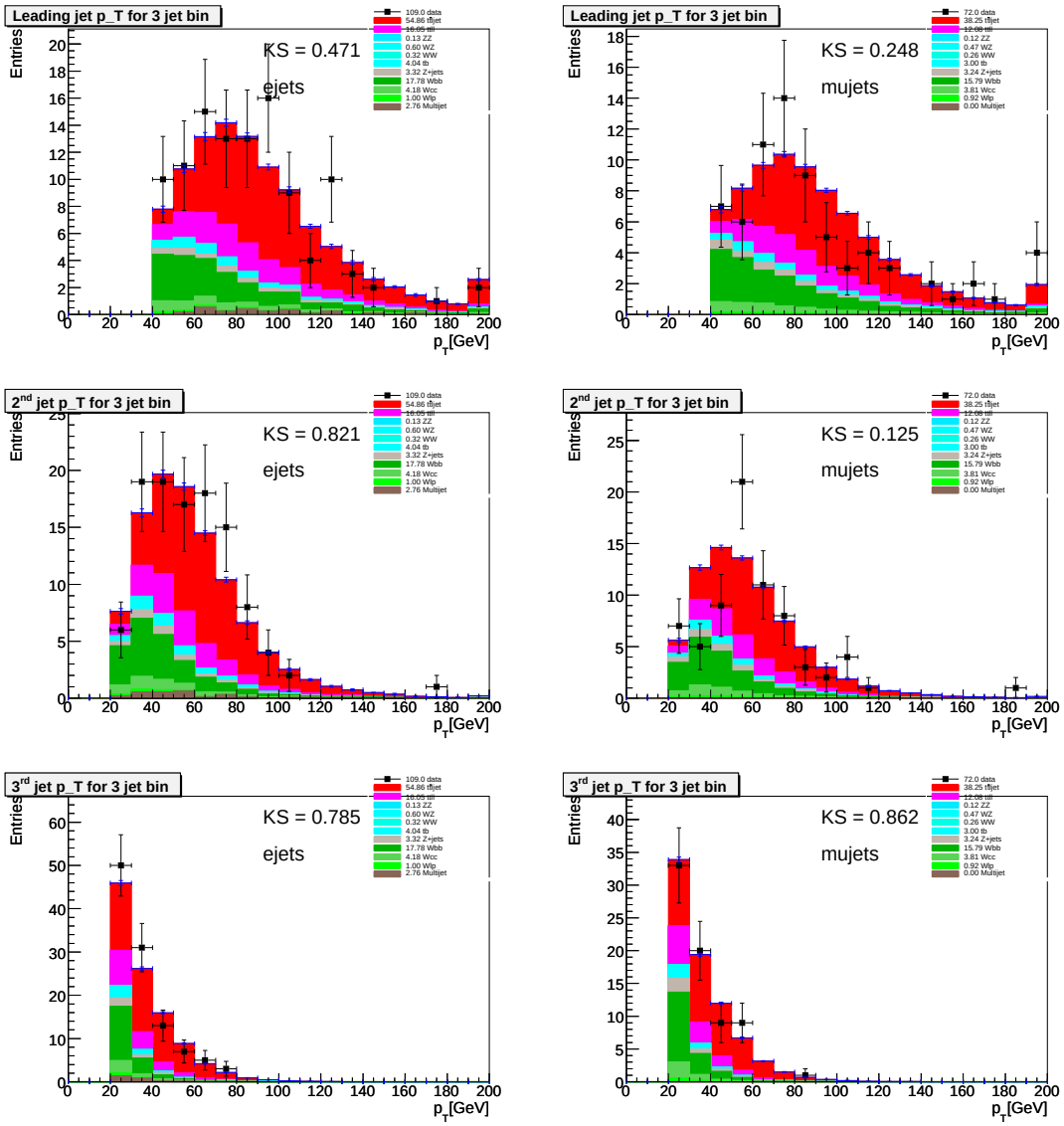


Figure C.46: Comparison for Run IIb between data and Monte Carlo for three jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet and p_T of the third leading jet.

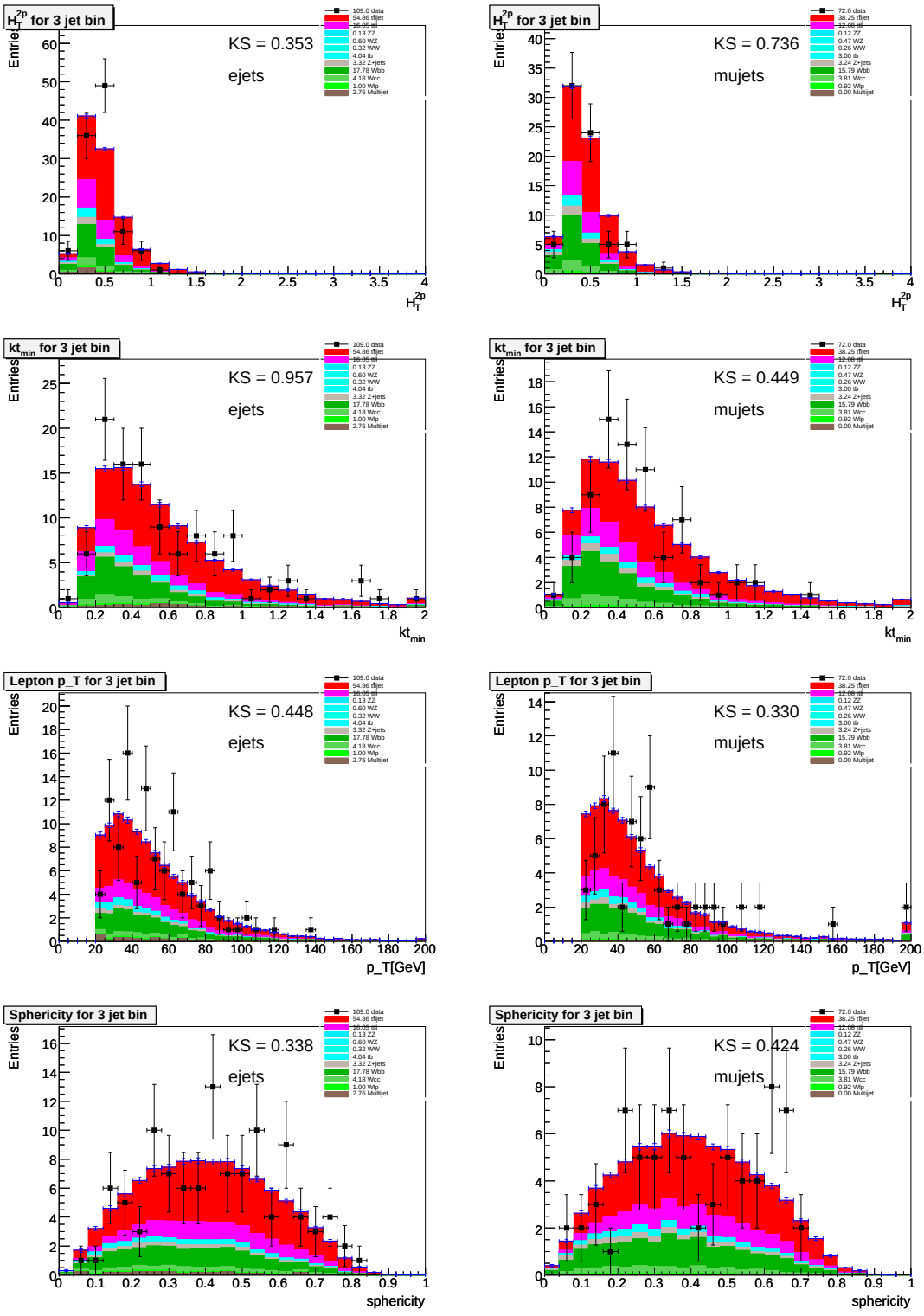


Figure C.47: Comparison for Run IIb between data and Monte Carlo for three jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

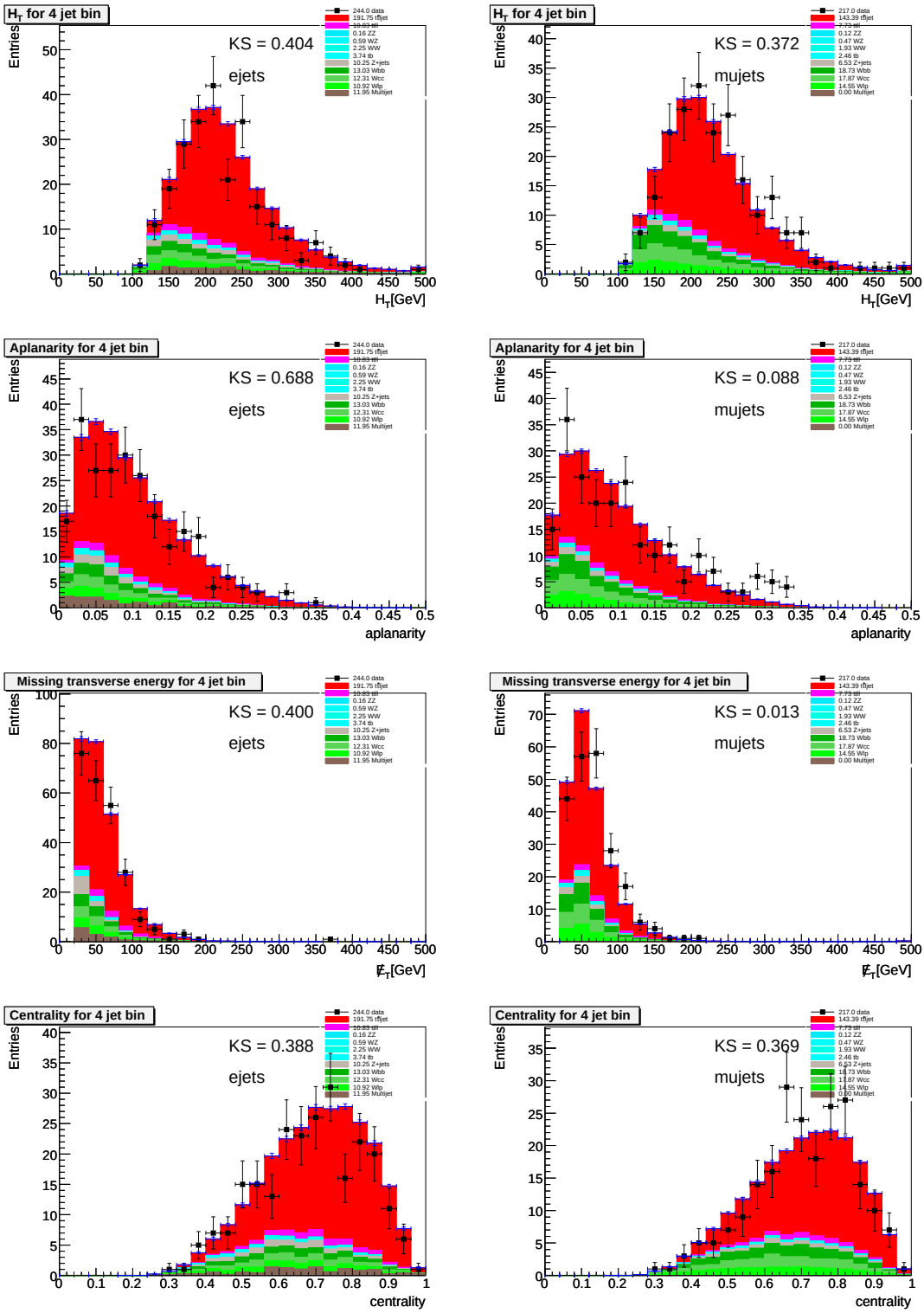


Figure C.48: Comparison for Run IIb between data and Monte Carlo for four or more jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

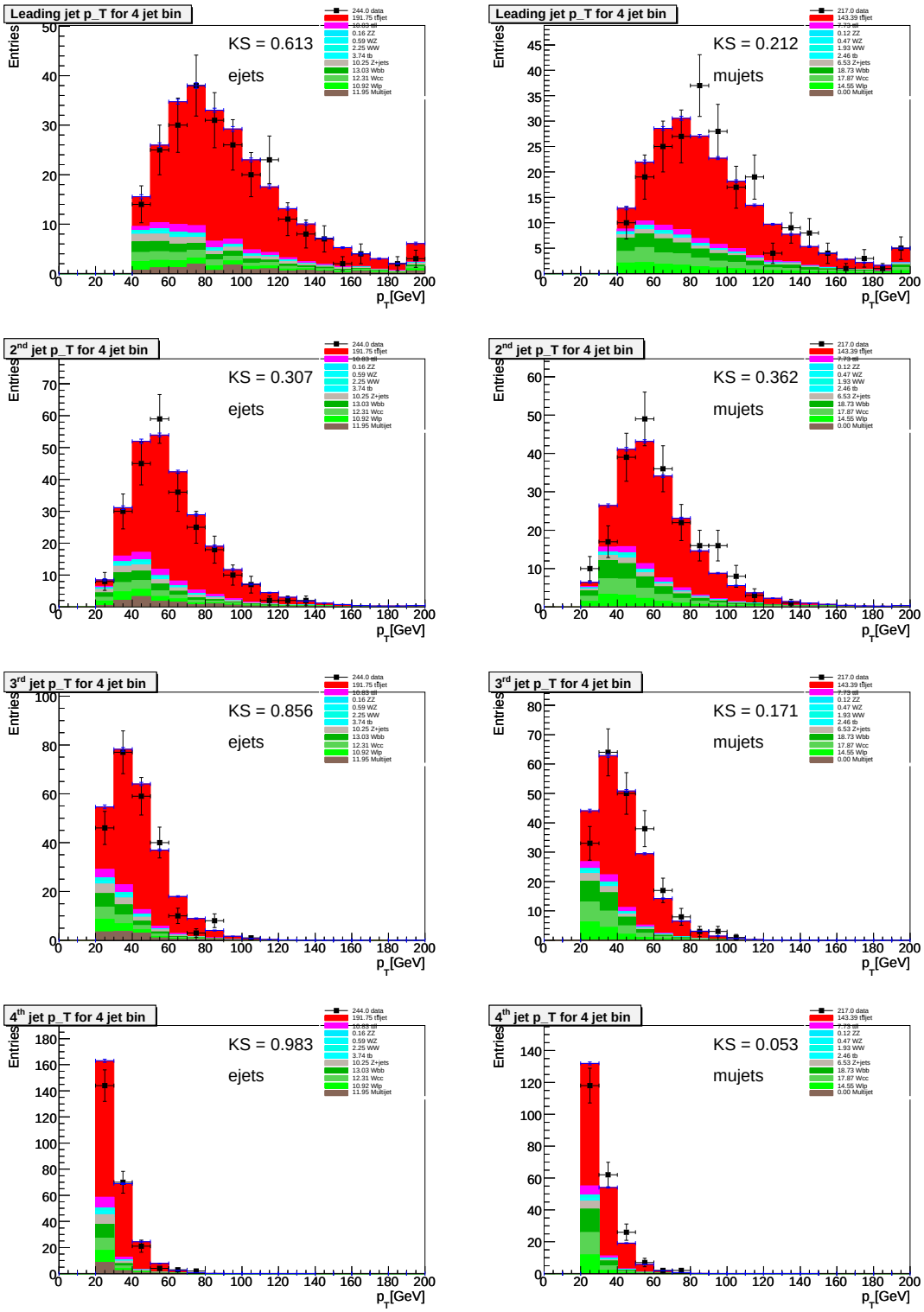


Figure C.49: Comparison for Run IIb between data and Monte Carlo for four or more jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

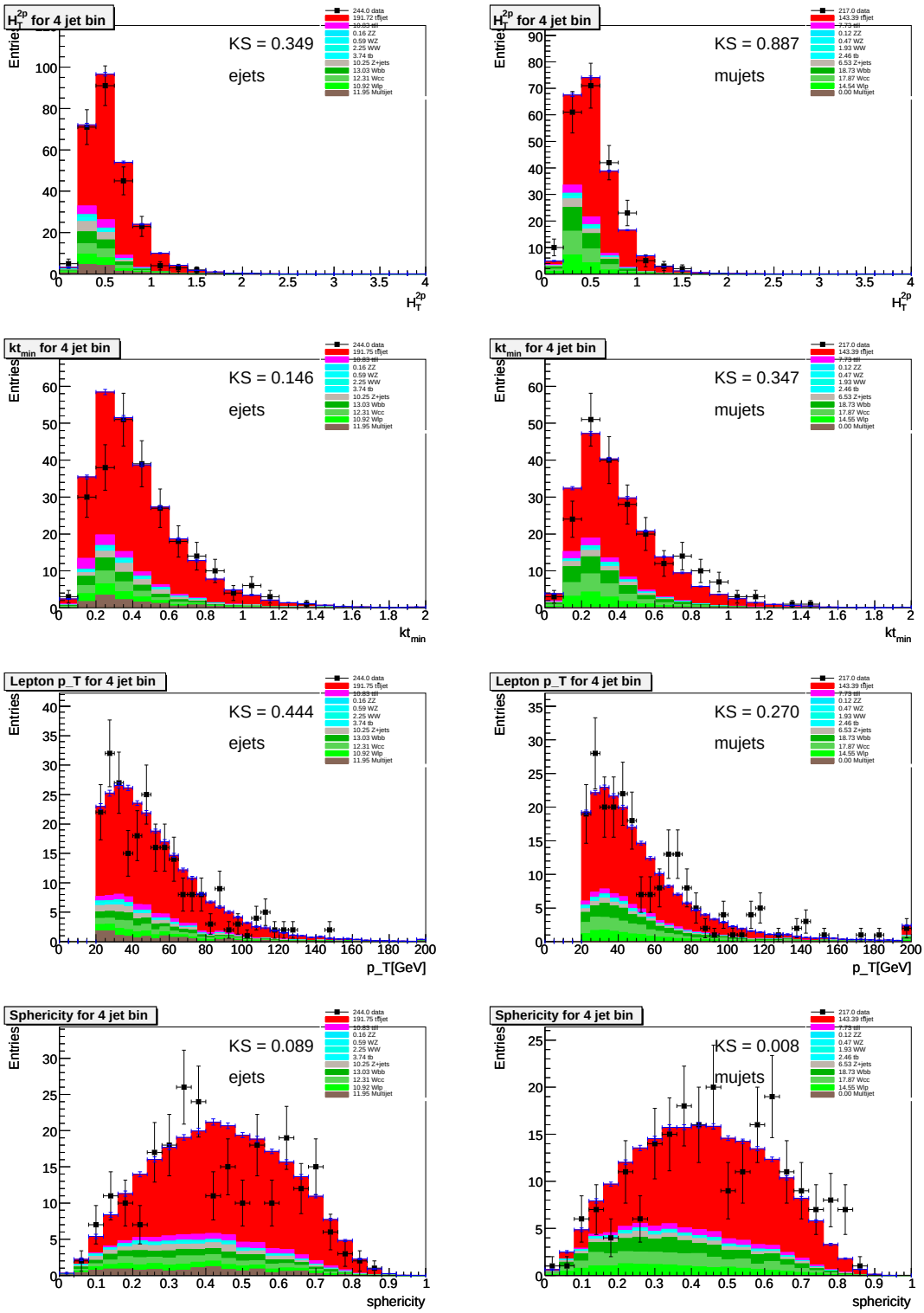


Figure C.50: Comparison for Run IIb between data and Monte Carlo for four or more jet events after all cuts including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

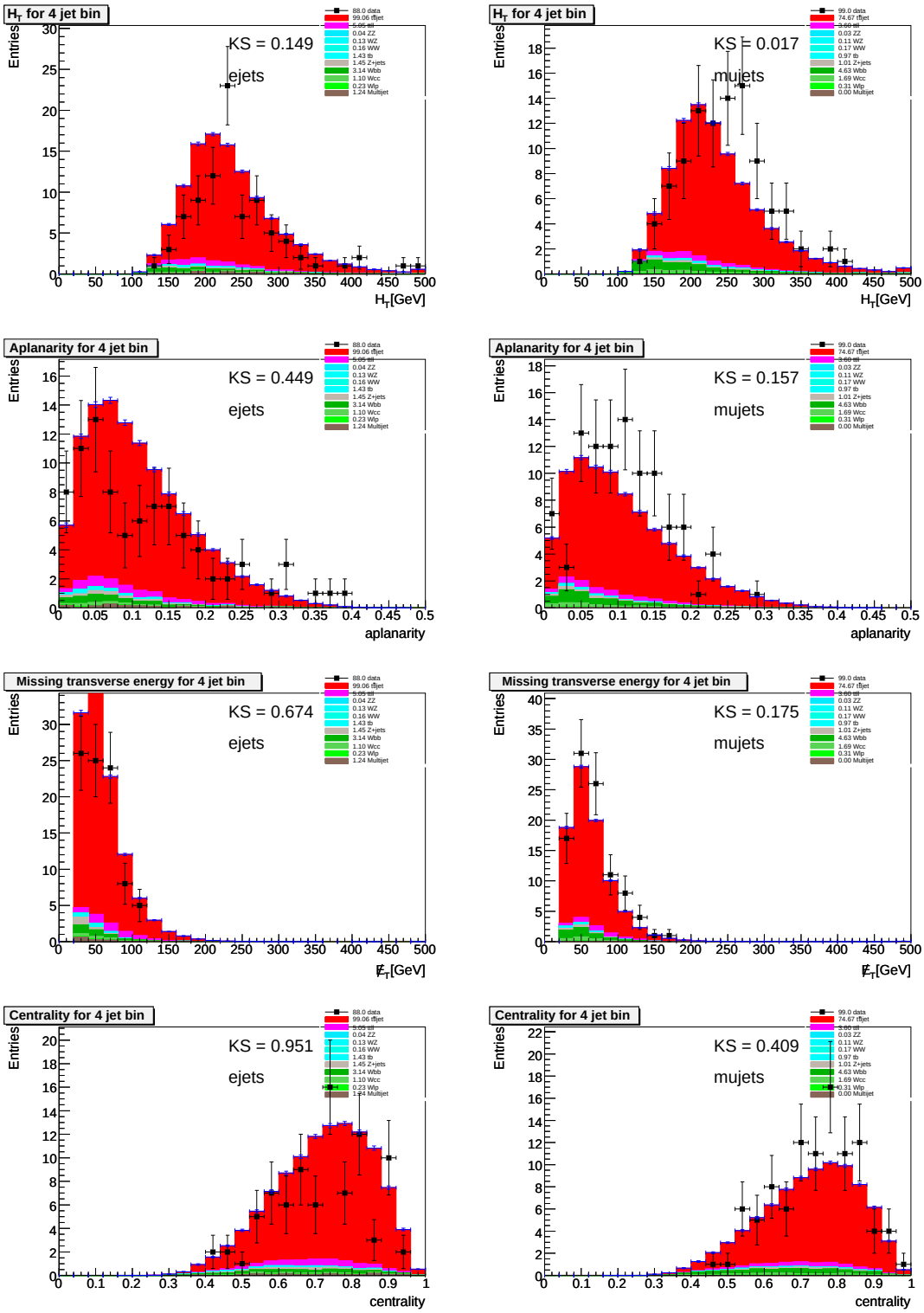


Figure C.51: Comparison for Run IIb between data and Monte Carlo for four or more jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

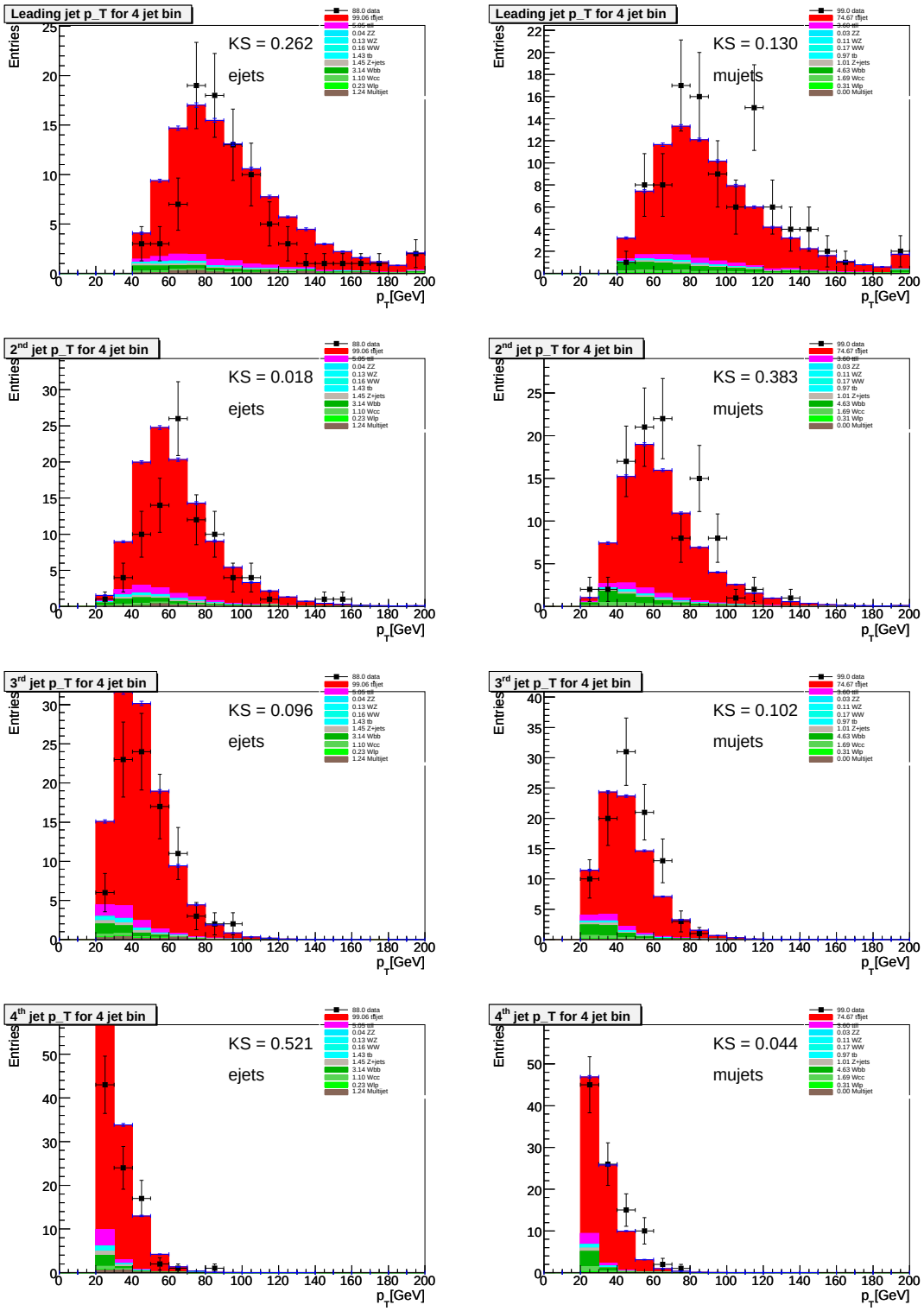


Figure C.52: Comparison for Run IIb between data and Monte Carlo for four or more jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet, p_T of the second leading jet, p_T of the third leading jet, and p_T of the fourth leading jet.

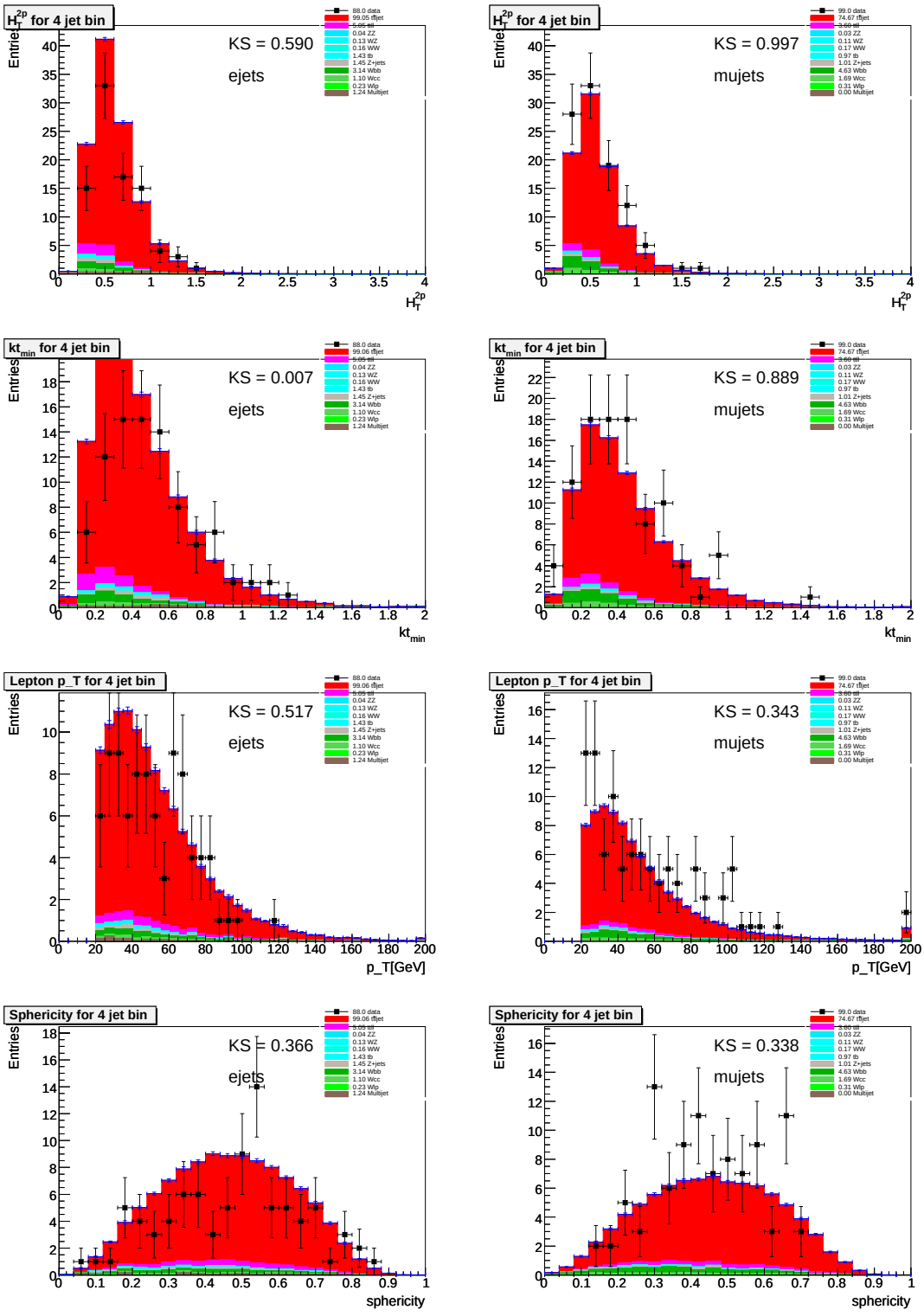


Figure C.53: Comparison for Run IIb between data and Monte Carlo for four or more jet events after all cuts including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

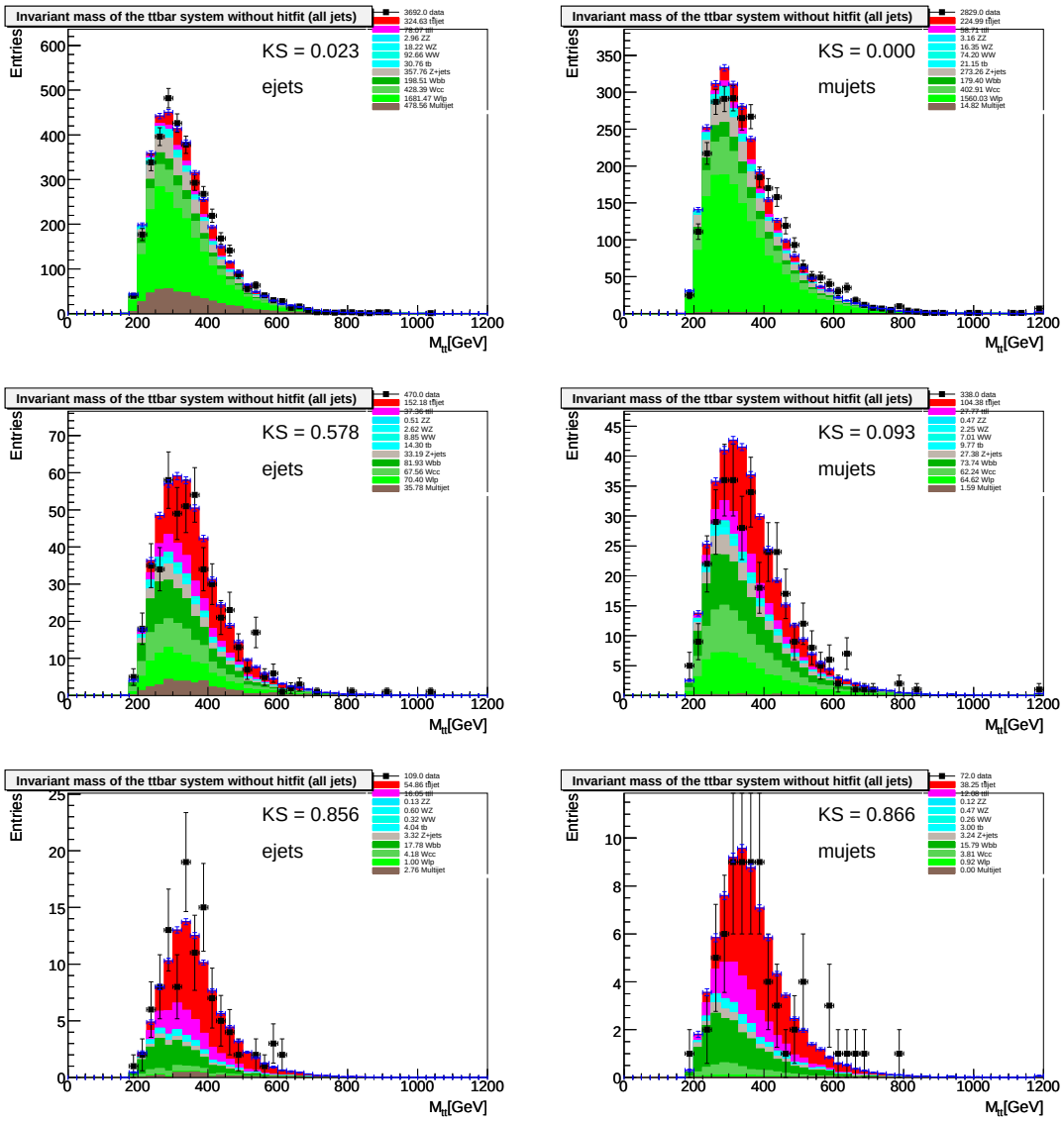


Figure C.54: Invariant mass distributions for three jet events in the e+jets (left), the μ+jets channel (right). First row corresponds to the selection before b-tagging. The second and third row show the results for $= 1$ and ≥ 2 medium NN b-tags, respectively.

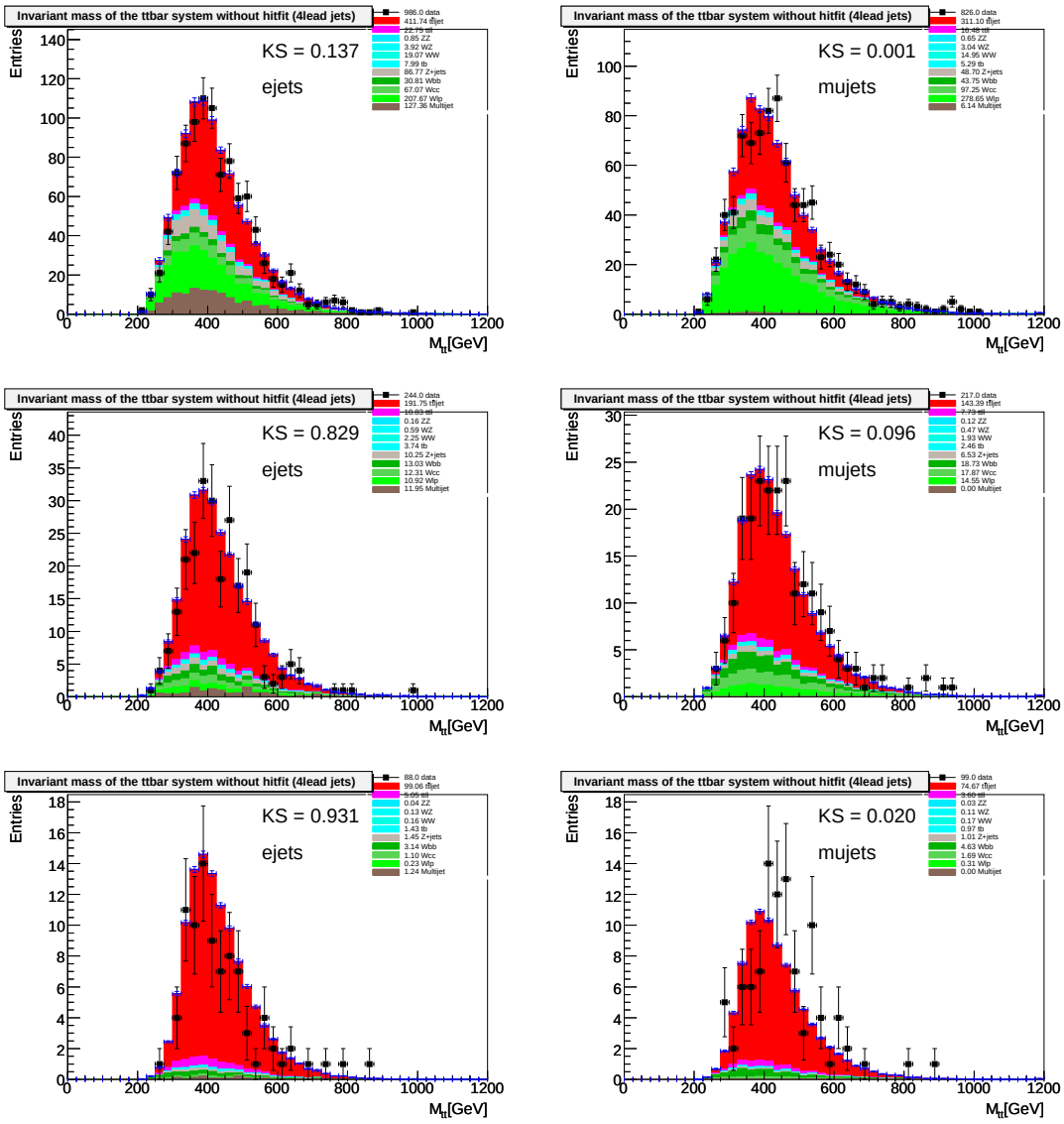


Figure C.55: Invariant mass distributions for four or more jet events in the e+jets (left), the μ+jets channel (right). First row corresponds to the selection before b-tagging. The second and third row show the results for $= 1$ and ≥ 2 medium NN b-tags, respectively.

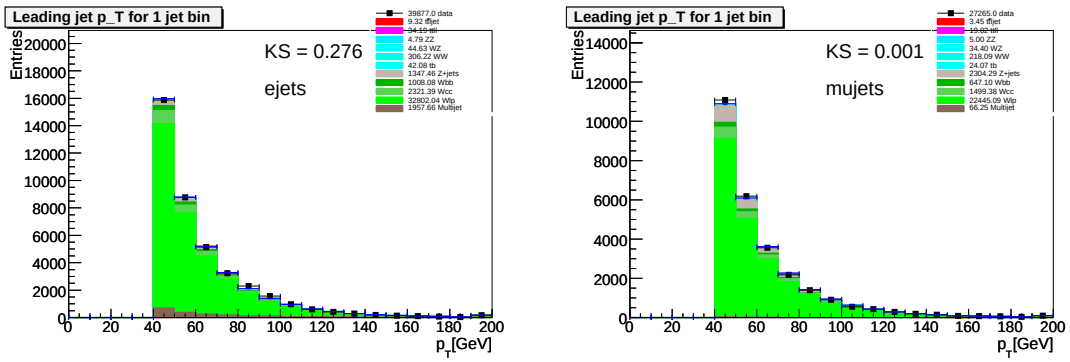


Figure C.56: Comparison for Run IIb between data and Monte Carlo with one jet before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables are: p_T of the leading jet.

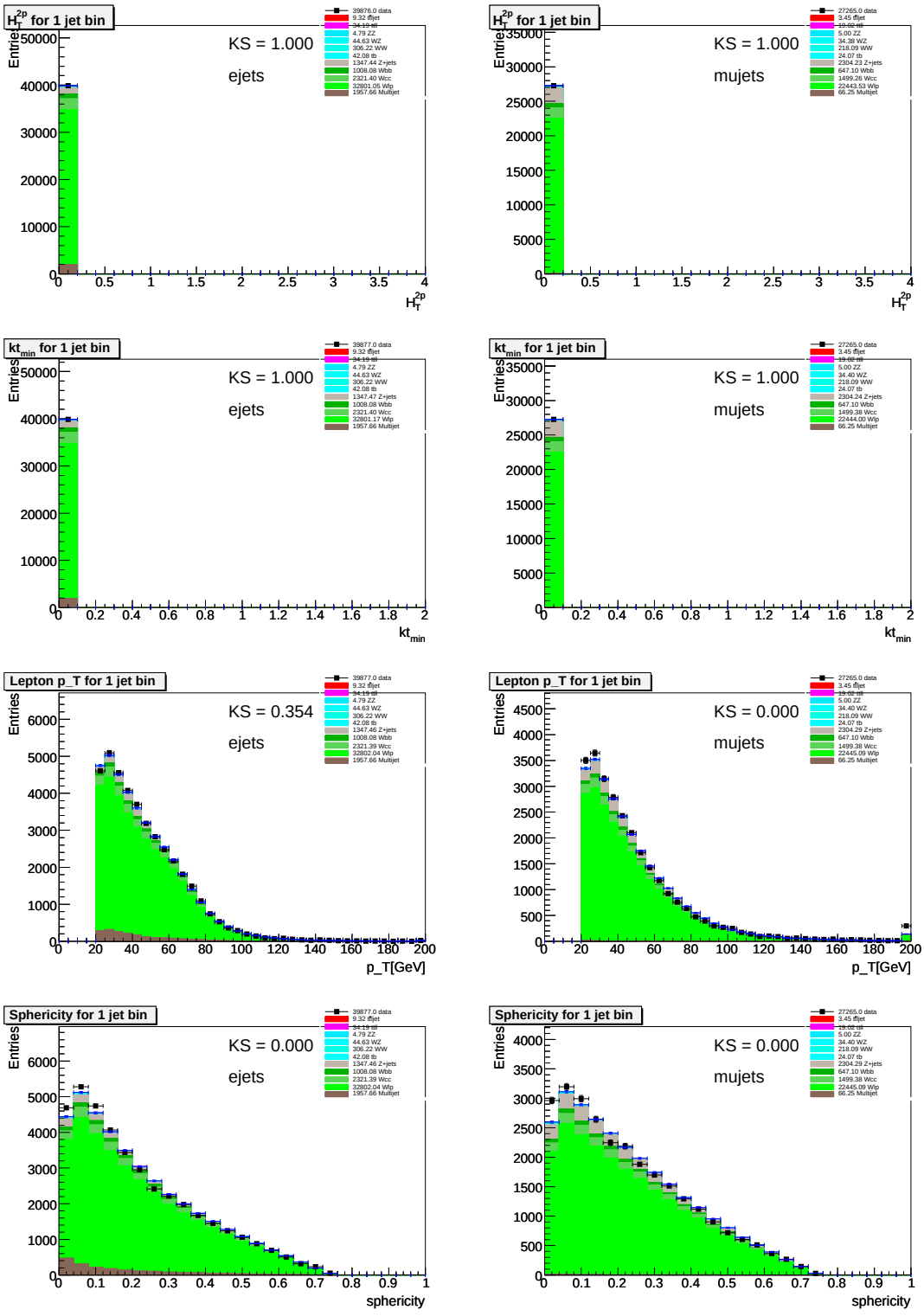


Figure C.57: Comparison for Run IIb between data and Monte Carlo with one jet before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

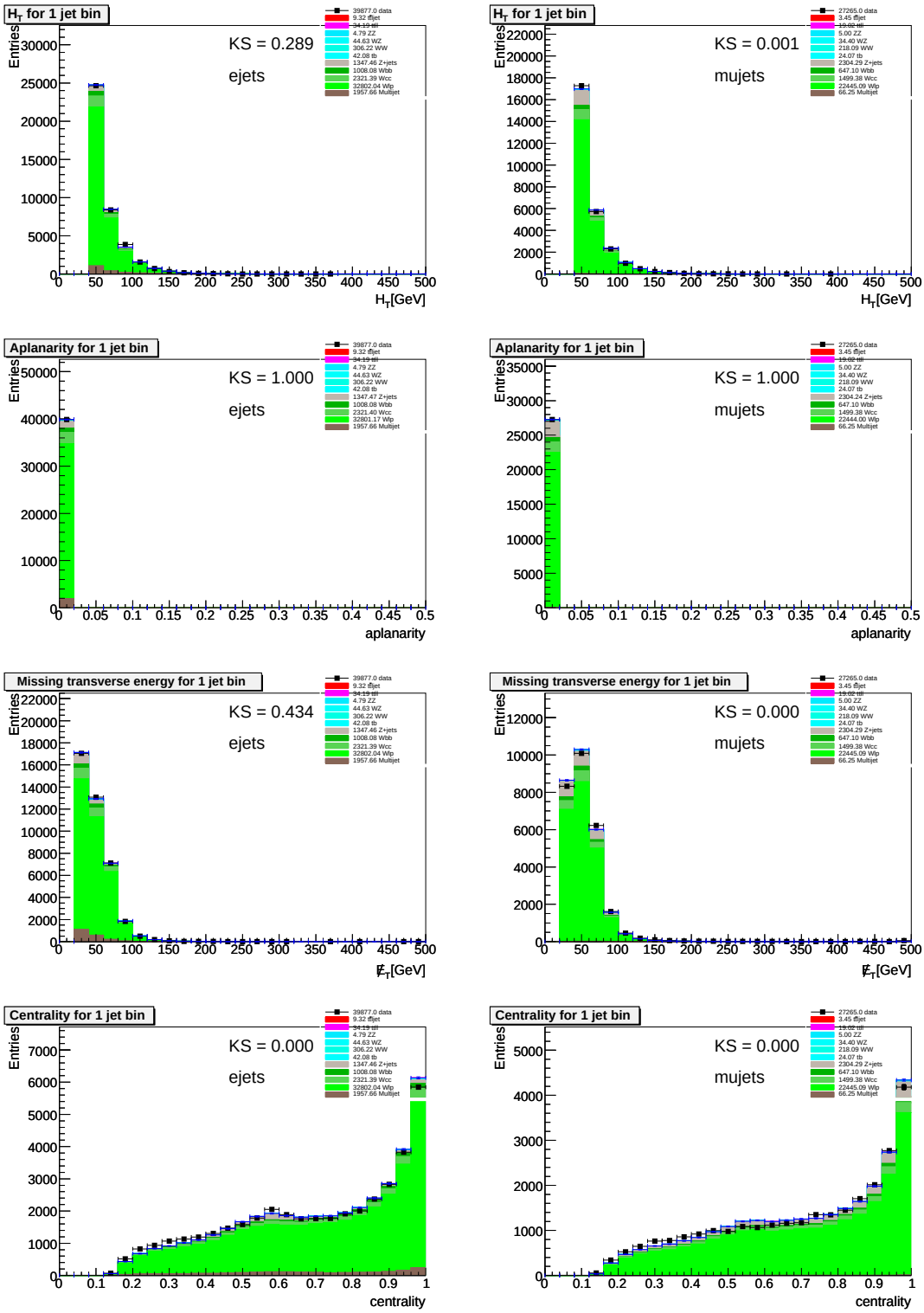


Figure C.58: Comparison for Run IIb between data and Monte Carlo with one jet before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

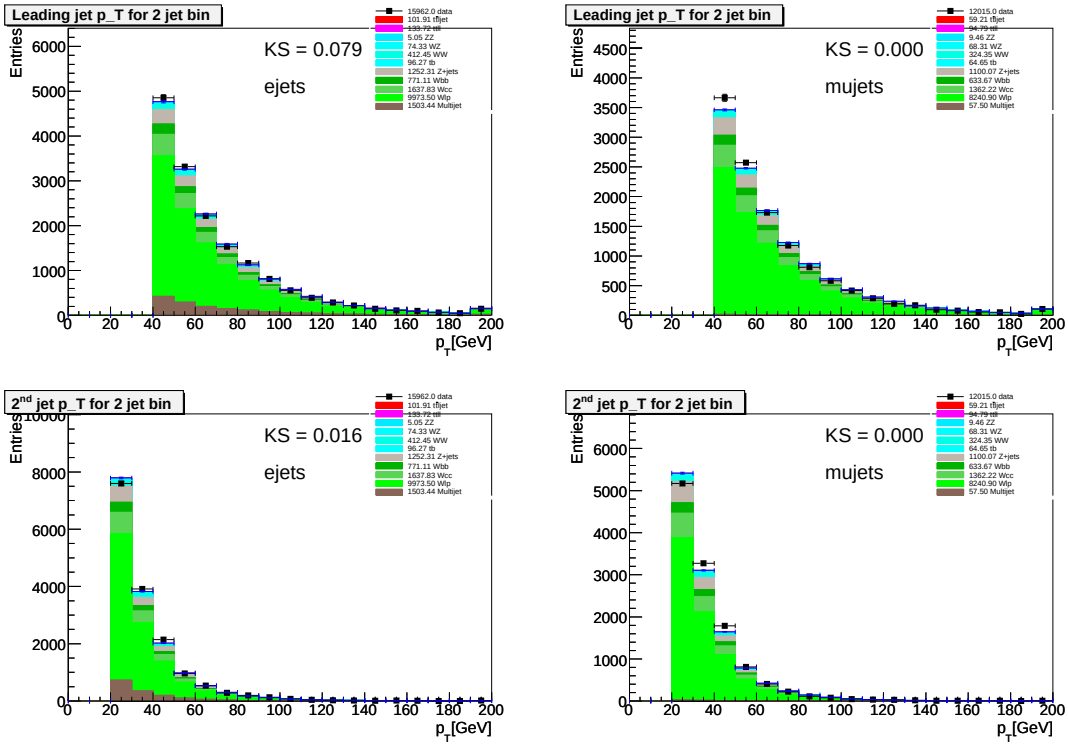


Figure C.59: Comparison for Run IIb between data and Monte Carlo with two jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet and p_T of the second leading jet.

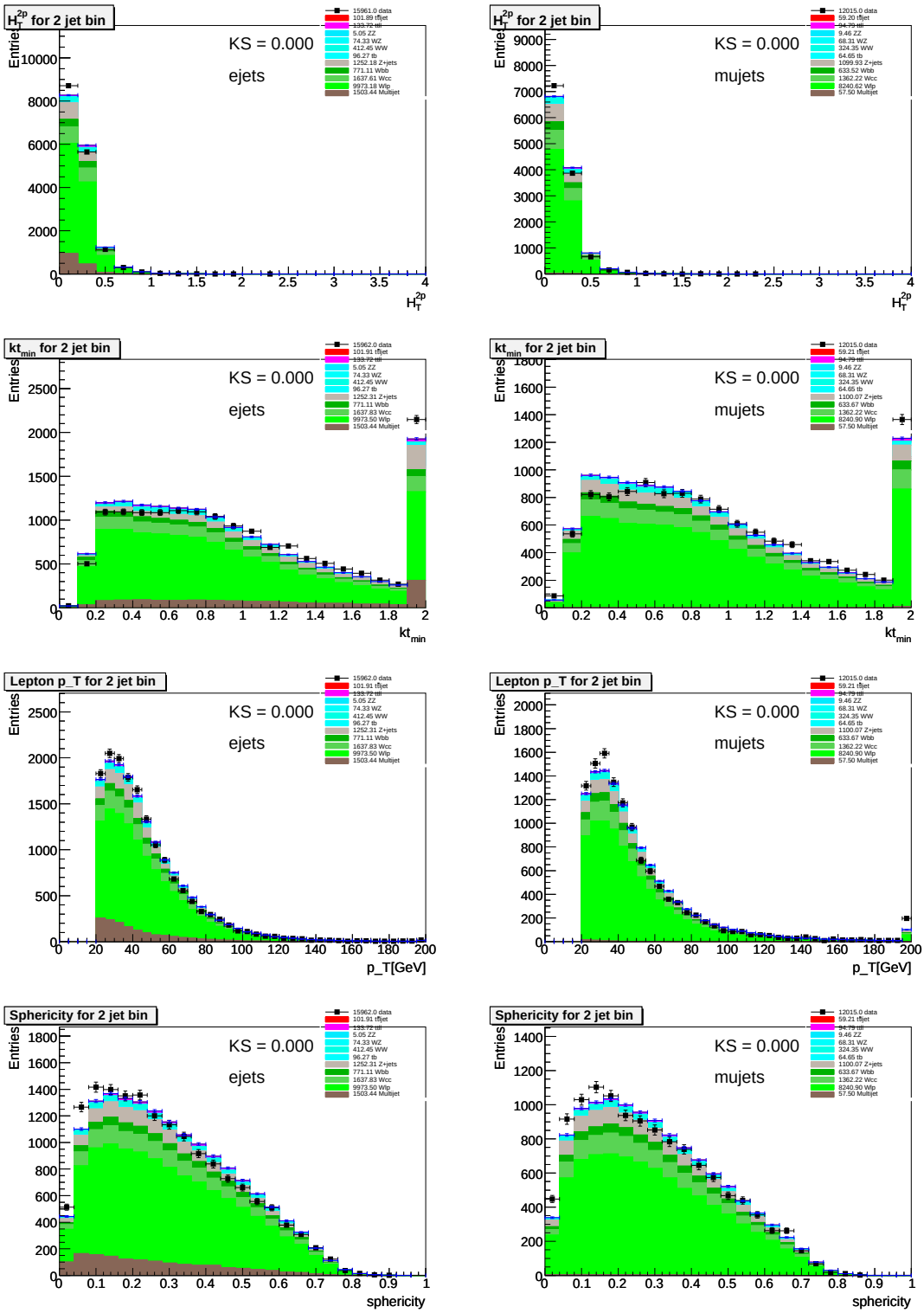


Figure C.60: Comparison for Run IIb between data and Monte Carlo with two jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

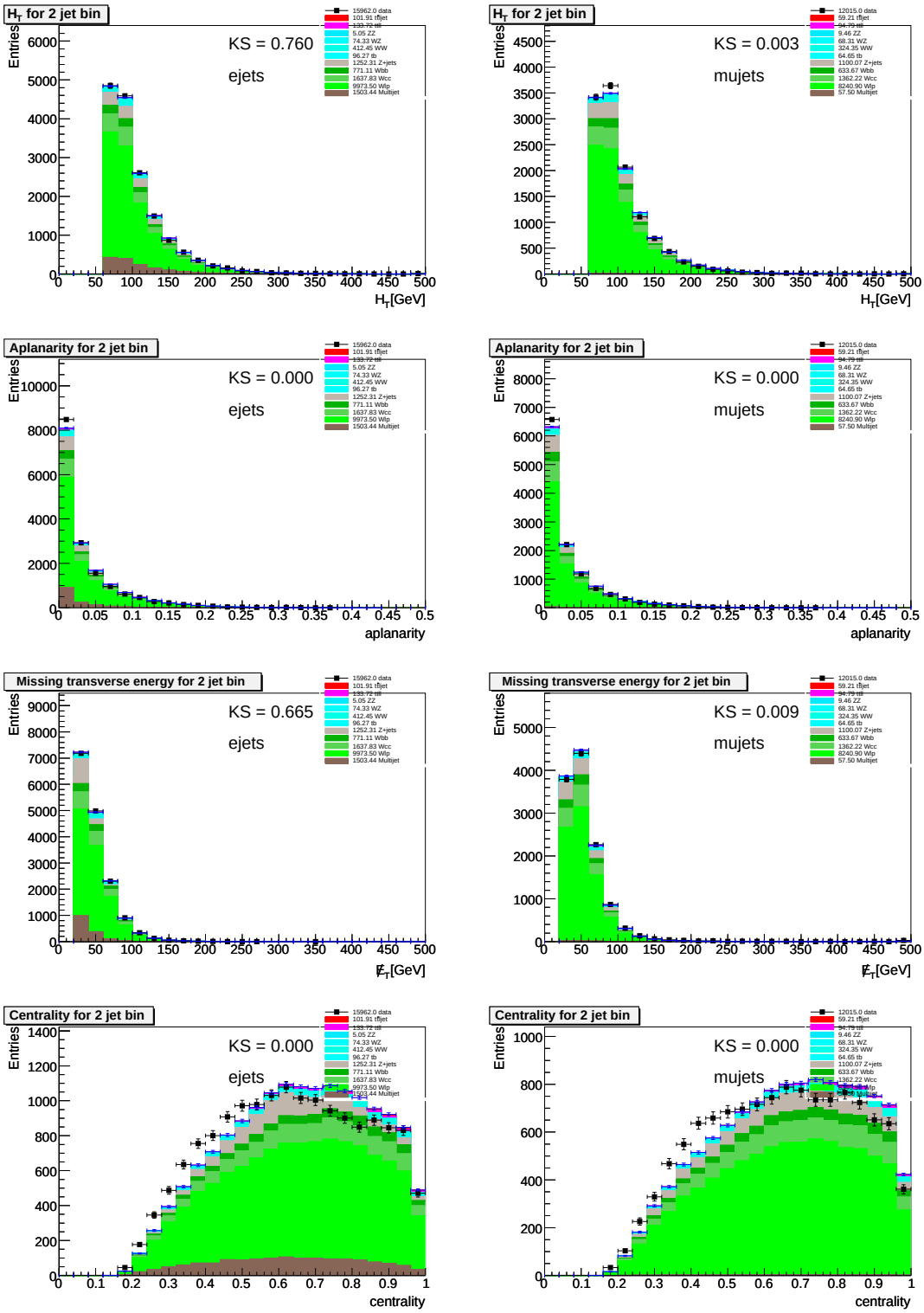


Figure C.61: Comparison for Run IIb between data and Monte Carlo with two jets before b-tagging in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

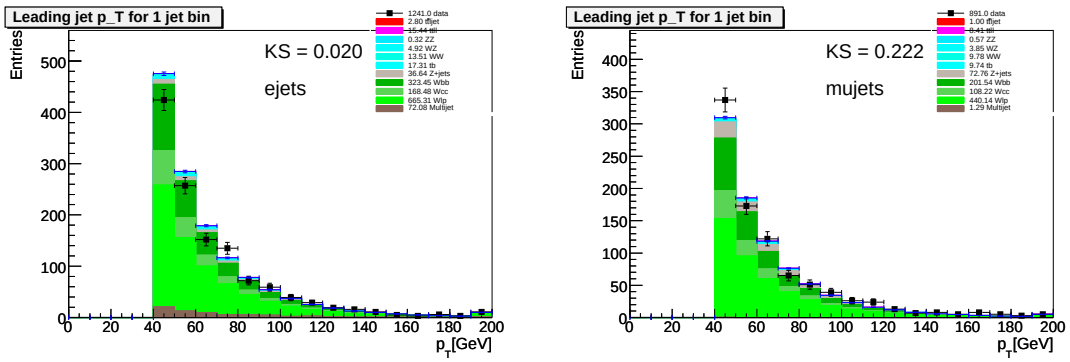


Figure C.62: Comparison for Run IIb between data and Monte Carlo with one jet including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables are: p_T of the leading jet.

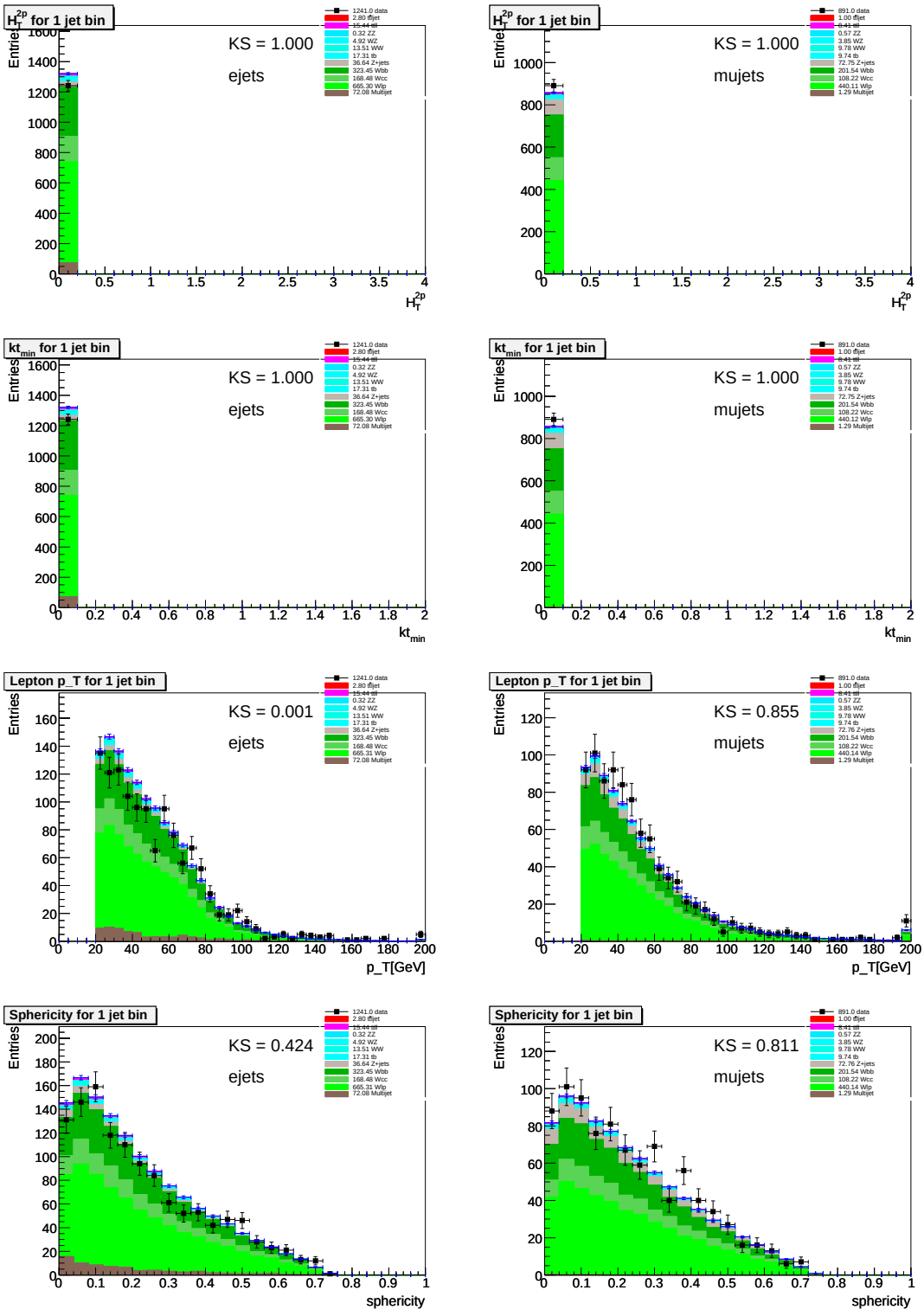


Figure C.63: Comparison for Run IIb between data and Monte Carlo with one jet including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

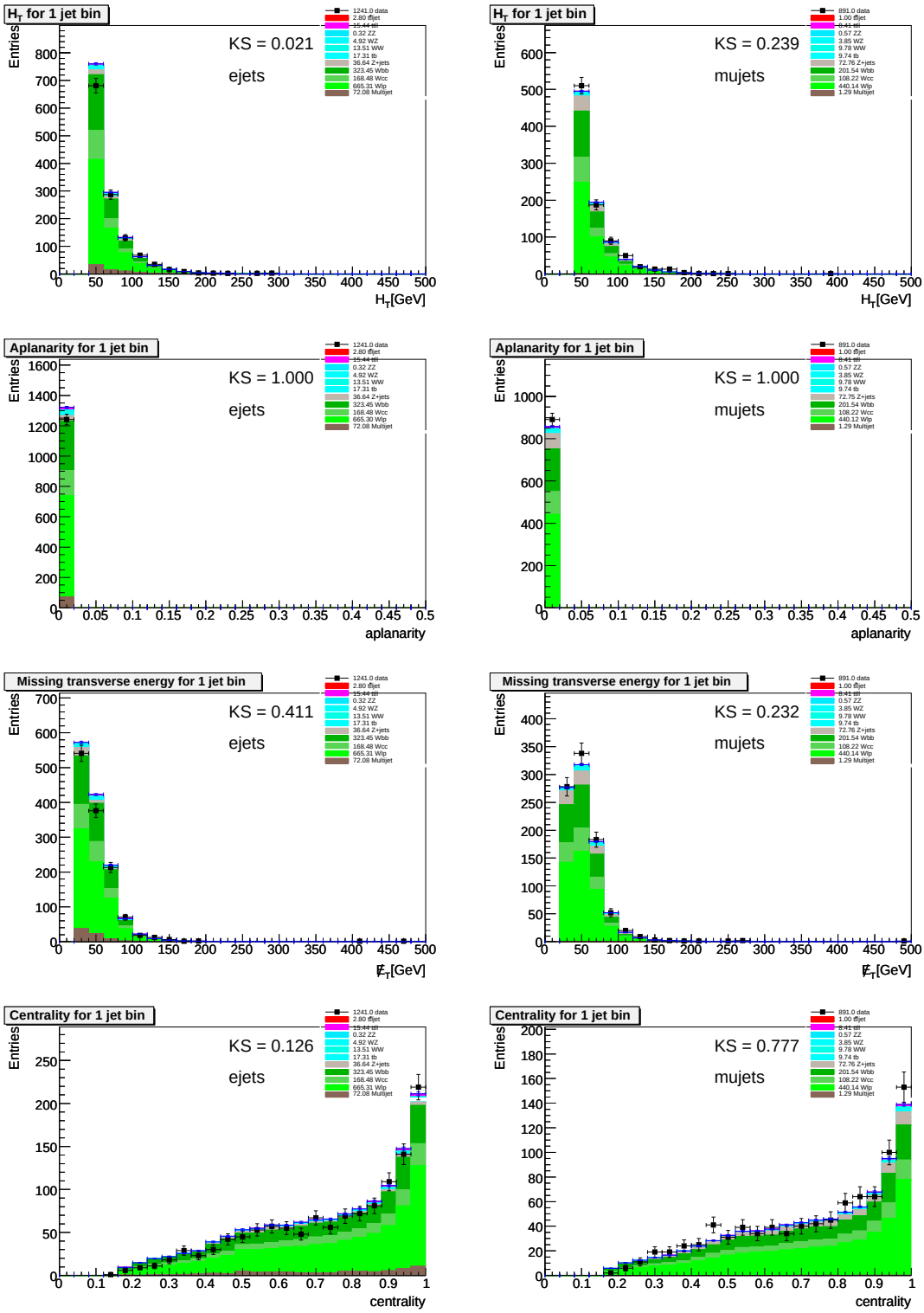


Figure C.64: Comparison for Run IIb between data and Monte Carlo with one jet including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

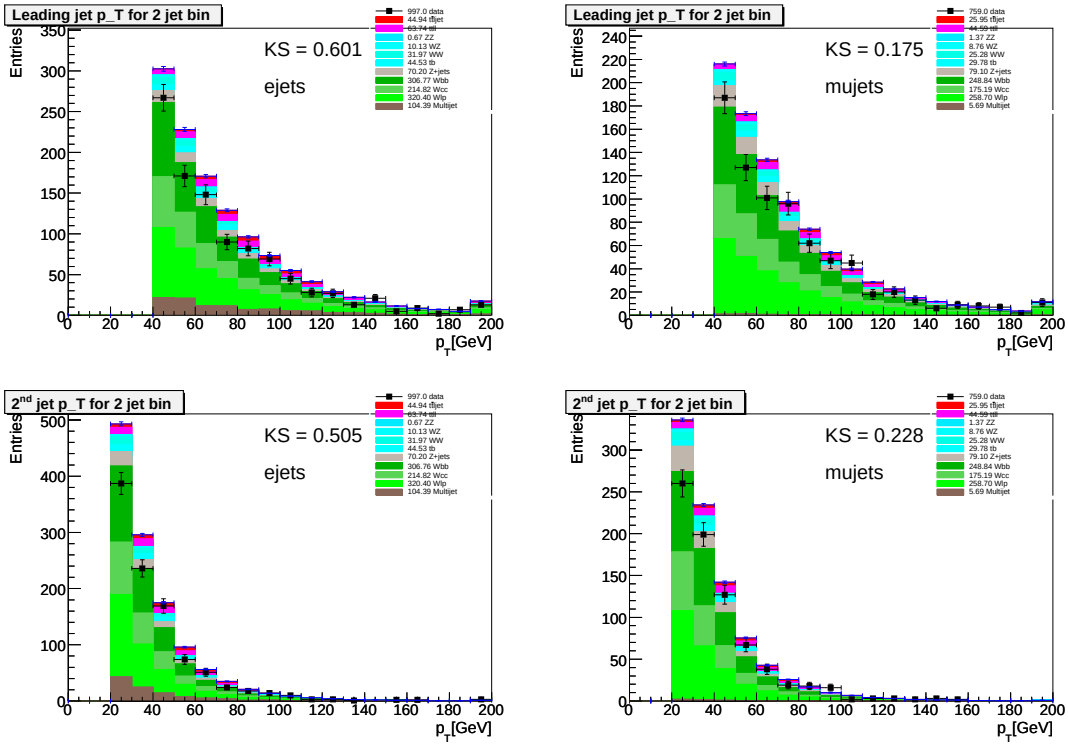


Figure C.65: Comparison for Run IIb between data and Monte Carlo with two jets including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet and p_T of the second leading jet.

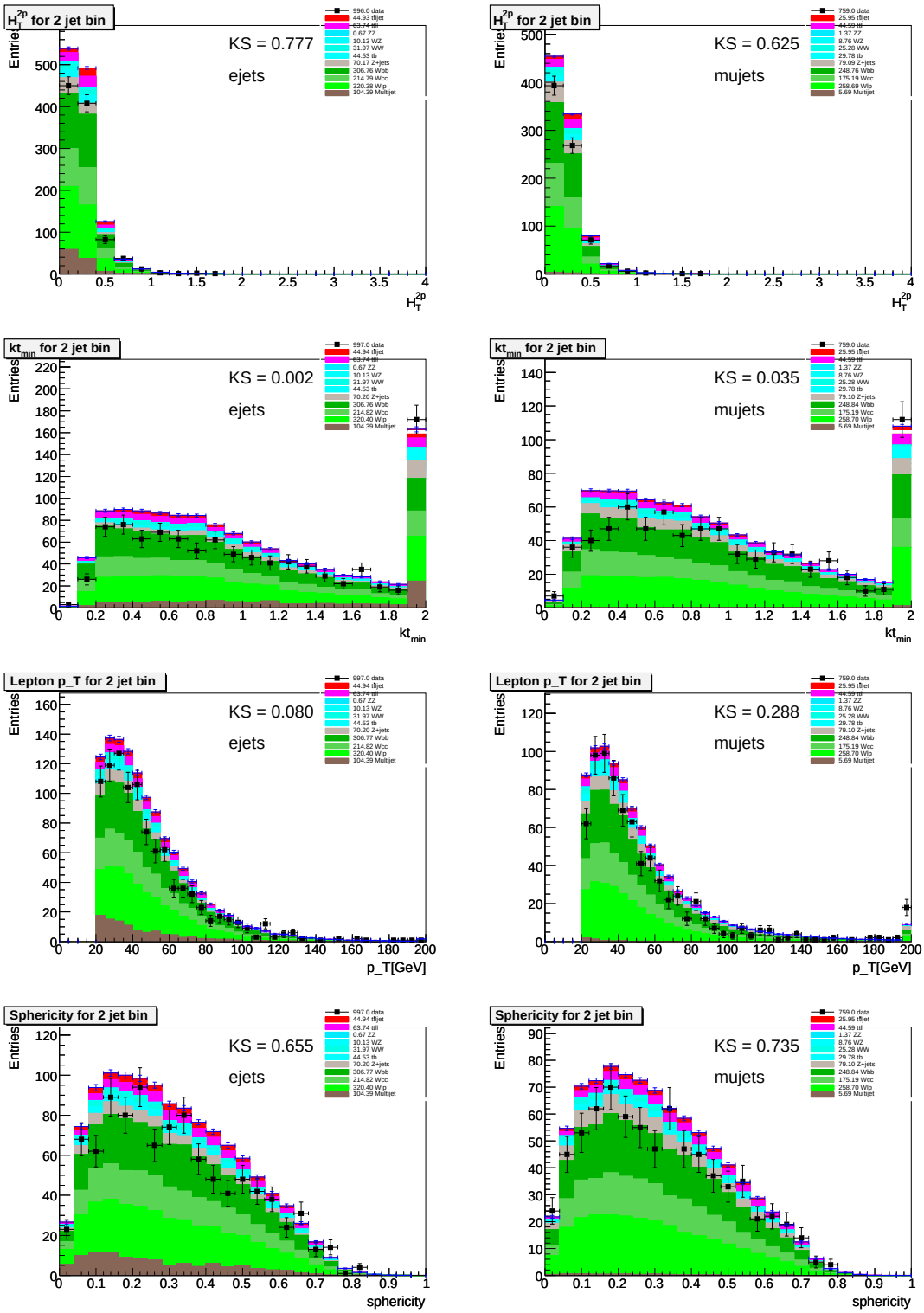


Figure C.66: Comparison for Run IIb between data and Monte Carlo with two jets including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

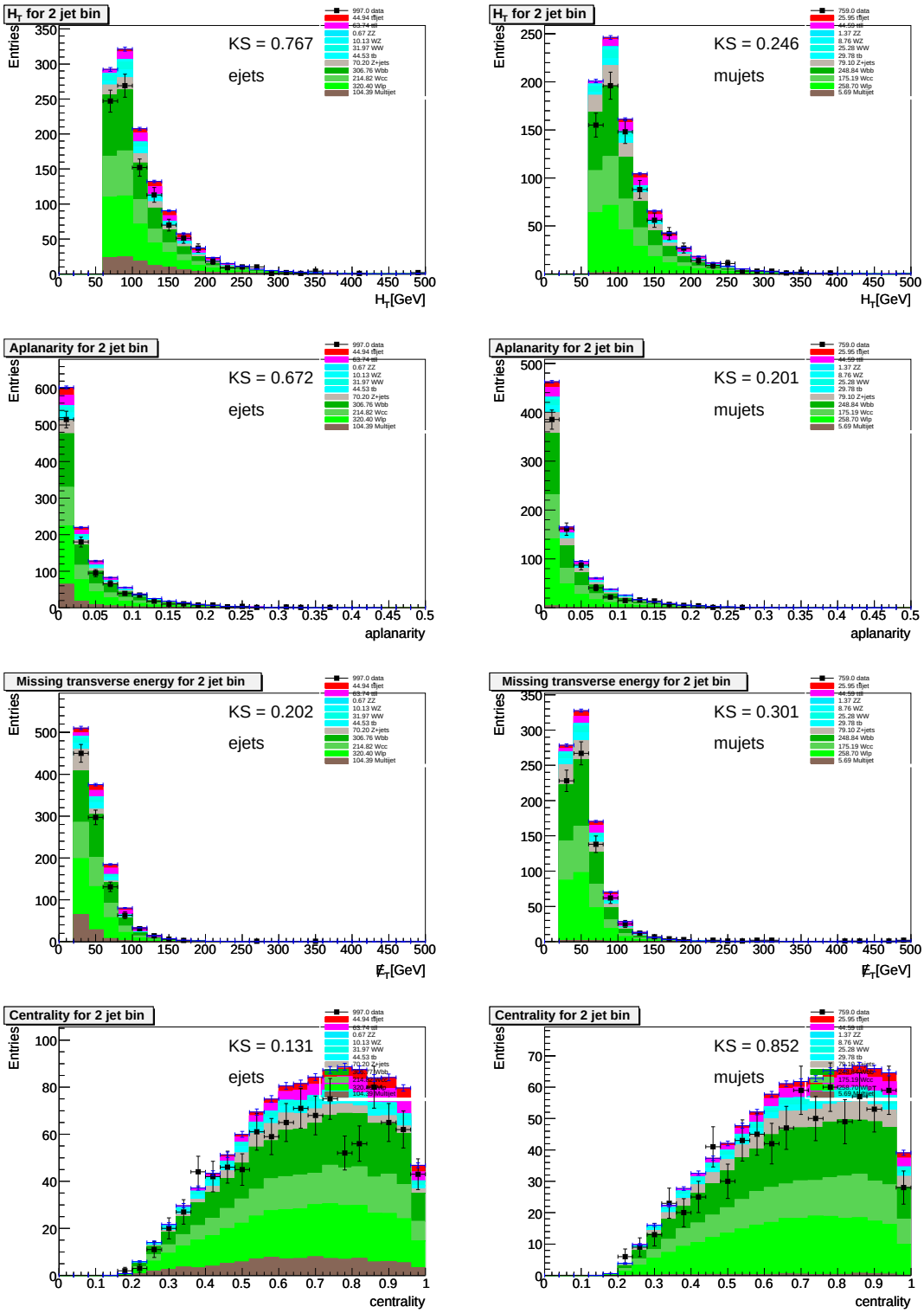


Figure C.67: Comparison for Run IIb between data and Monte Carlo with two jets including one medium NN b-tag in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

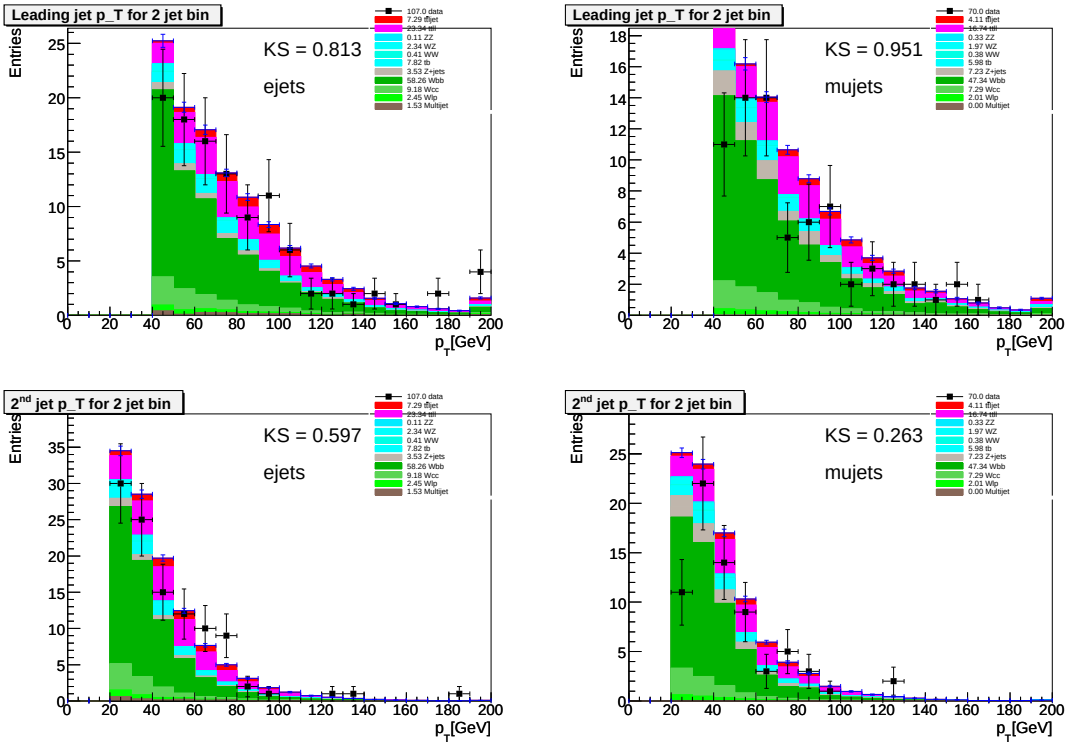


Figure C.68: Comparison for Run IIb between data and Monte Carlo with two jets including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T of the leading jet and p_T of the second leading jet.

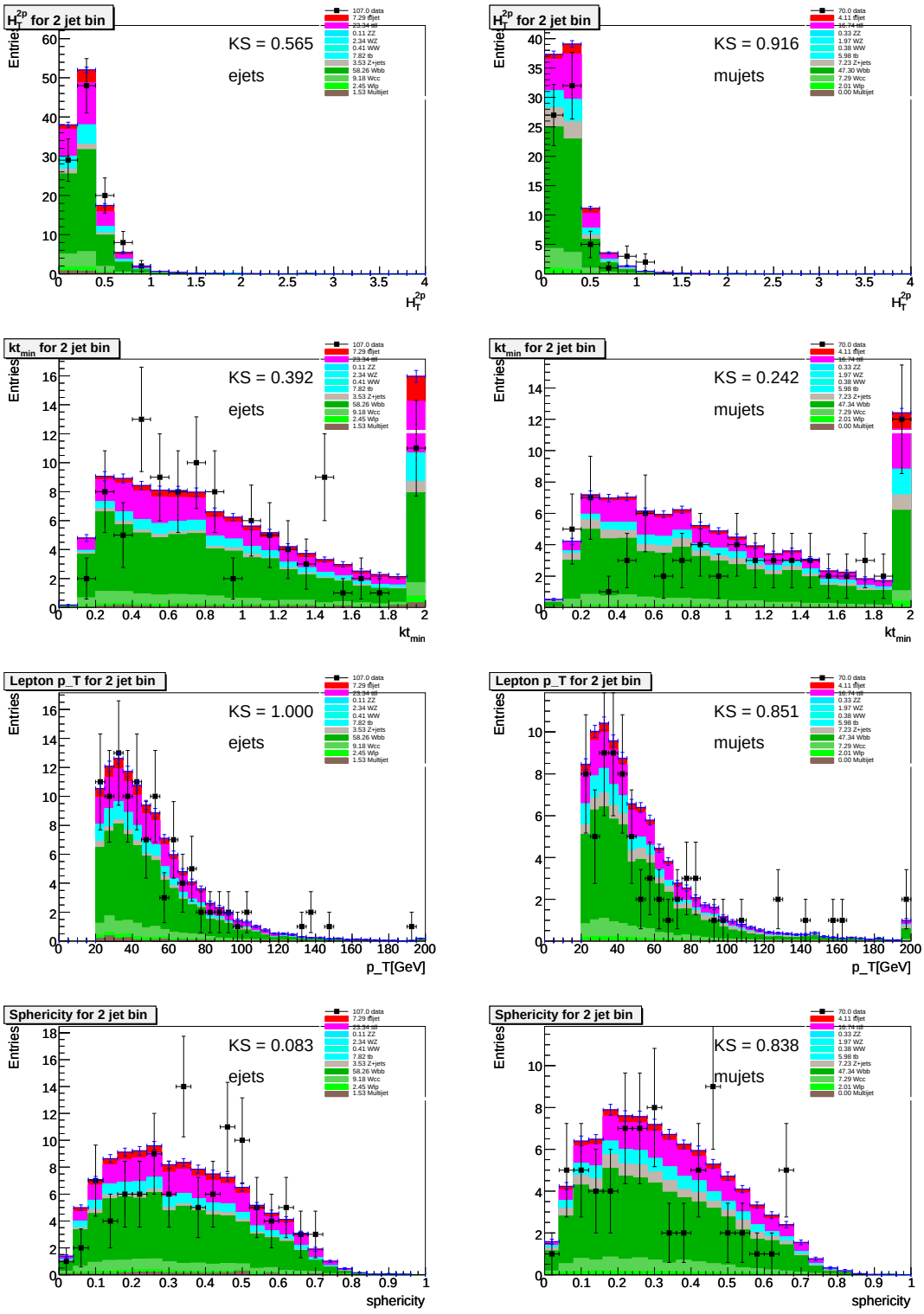


Figure C.69: Comparison for Run IIb between data and Monte Carlo with two jets including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: ht2p, Ktminp, the p_T of the lepton, and sphericity.

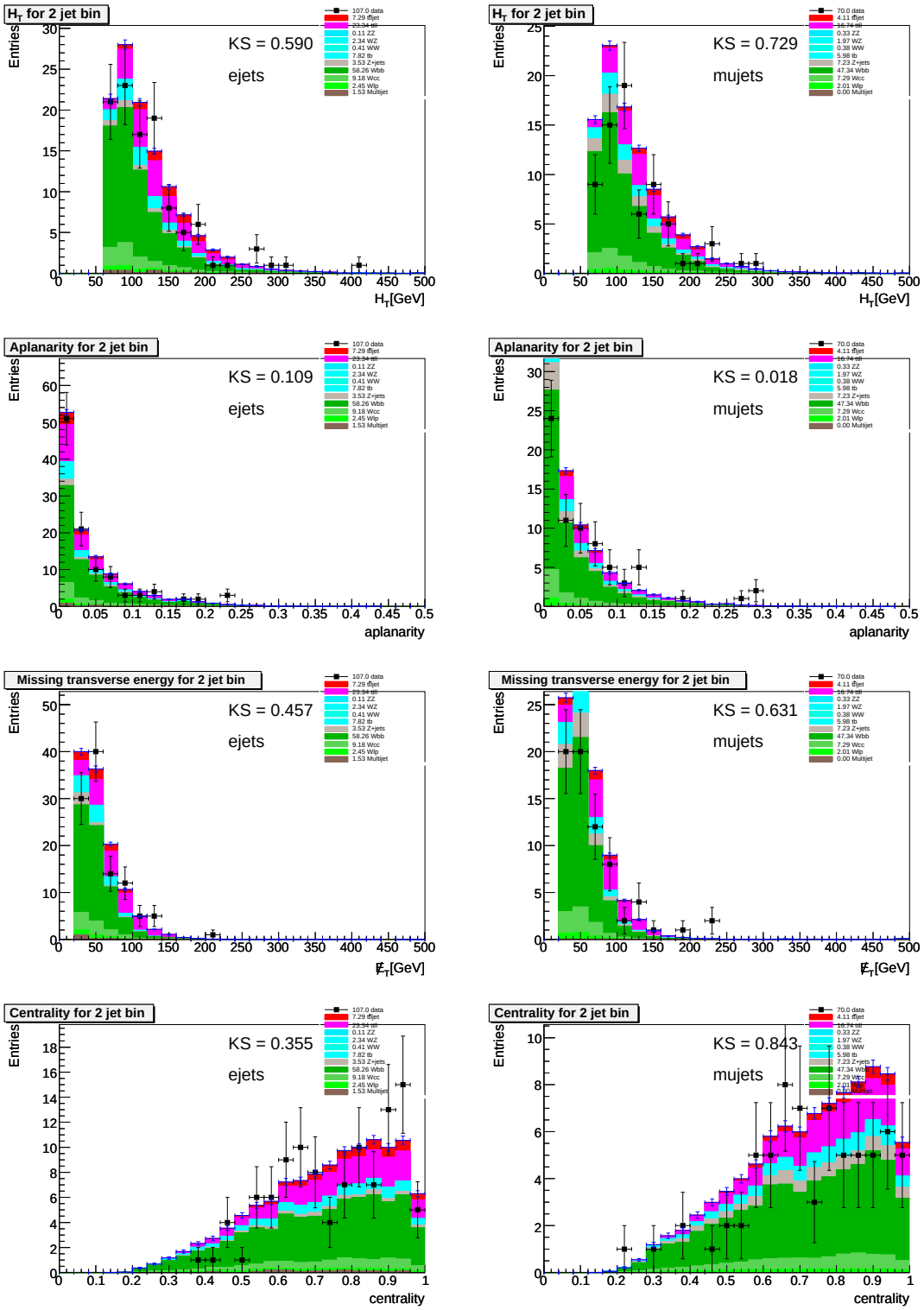


Figure C.70: Comparison for Run IIb between data and Monte Carlo with two jets including two or more medium NN b-tags in the e+jets channel (left) and the μ+jets channel (right). The variables from top to bottom are: p_T sum of all jets (Ht), aplanarity, missing E_T, and centrality.

List of Figures

List of Tables

Bibliography

[1] F. Halzen and A. D. Martin, *Quarks and leptons: An introductory course in modern particle physics*, Wiley, 1984.

[2] R. K. Ellis, W. J. Stirling, and B. R. Webber, *Qcd and collider physics*, Cambridge Univ. Press, 2003.

[3] D. Griffiths, *Introduction to elementary particles*, John Wiley & Sons, New York, USA, 1987.

[4] D. Perkins, *Introduction to high energy physics*, Addison-Wesley, Reading, USA, 1982.

[5] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964), 168–171.

[6] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967), 1264–1266.

[7] V. M. Abazov et al., *Evidence for an anomalous like-sign dimuon charge asymmetry*, (2010), 1005.2757.

[8] A. Leike, *The phenomenology of extra neutral gauge bosons*, Phys. Rept. **317** (1999), 143, hep-ph/9805494.

[9] B. Lillie, L. Randall, and L.-T. Wang, *The bulk RS KK-gluon at the LHC*, JHEP **09** (2007), 074, hep-ph/0701166.

[10] T. G. Rizzo, *Testing the nature of Kaluza-Klein excitations at future lepton colliders*, Phys. Rev. **D61** (2000), 055005, hep-ph/9909232.

[11] L. M. Sehgal and M. Wanninger, *Forward - backward asymmetry in two jet events: Signature of axigluons in proton - anti-proton collisions*, Phys. Lett. **B 200** (1988), 211.

[12] C. T. Hill and S. Parke, *Top production: sensitivity to new physics*, Phys. Rev. **D 49** (1994), 4454.

[13] V. M. Abazov et al., *Simultaneous measurement of the ratio $B(t \to Wb)$ $/B(t \to Wq)$ and the top quark pair production cross section with the D0 detector at $\sqrt{s} = 1.96$-TeV*, Phys. Rev. Lett. **100** (2008), 192003, 0801.1326.

[14] R. M. Harris, C. T. Hill, and S. J. Parke, *Cross section for topcolor $Z'(t)$ decaying to t anti-t*, (1999), hep-ph/9911288.

[15] V. M. Abazov et al., *Search for Narrow $t\bar{t}$ Resonances in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **92** (2004), 221801.

[16] T. Affolder et al., *Search for new particles decaying to $t\bar{t}$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **85** (2000), 2062.

[17] V. M. Abazov et al., *Search for ttbar resonances in the lepton plus jets final state in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Lett. **B668** (2008), 98–104, 0804.3664.

[18] V. M. Abazov et al., *Search for $t\bar{t}$ resonances in the lepton+jets final state in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, (2008), DØ Note 5600 conf.

[19] T. Aaltonen et al., *Limits on the production of narrow $t\bar{t}$ resonances in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. **D77** (2008), 051102, 0710.5335.

[20] T. Aaltonen et al., *Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **100** (2008), 231801, 0709.0705.

[21] T. Aaltonen et al., *Massive gluon search in the top pair lepton+jets channel using Dynamical Likelihood Method with $1.9\,fb^{-1}$*, (2008), CDF Note 9164.

[22] T. Aaltonen et al., *Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV*, Phys. Rev. Lett. **100** (2008), 231801, 0709.0705.

[23] J. Goldstone, A. Salam, and S. Weinberg, *Broken symmetries*, Phys. Rev. **127** (1962), 965–970.

[24] C. Amsler et al., *Review of particle physics*, Phys. Lett. **B667** (2008), 1.

[25] Tevatron Electroweak Working Group, *Combination of CDF and D0 Results on the Mass of the Top Quark*, (2009), 0903.2503.

[26] M. Gell-Mann, *Symmetries of baryons and mesons*, Phys. Rev. **125** (1962), 1067–1084.

[27] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, Nucl. Phys. **26** (1961), 222–229.

[28] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. 2*, (1964), "CERN-TH-412".

[29] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961), 579–588.

[30] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964), 132–133.

[31] M. Kobayashi and T. Maskawa, CP *violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973), 652–657.

[32] Z. Maki, M. Nakagawa, and S. Sakata, (1962), Prog. Theor. Phys. **28**, 870.

[33] S. Abachi et al., *Observation of the top quark*, Phys. Rev. Lett. **74** (1995), 2632–2637, hep-ex/9503003.

[34] F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995), 2626–2631, hep-ex/9503002.

[35] W. Wagner, *Top quark physics in hadron collisions*, Rept. Prog. Phys. **68** (2005), 2409–2494, hep-ph/0507207.

[36] D. Wicke, *Properties of the Top Quark*, (2009), 1005.2460.

[37] S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. D **78** (2008), 0804.1476.

[38] `http://durpdg.dur.ac.uk/hepdata/pdf3.html`.

[39] V. M. Abazov et al., *Observation of Single Top-Quark Production*, Phys. Rev. Lett. **103** (2009), 092001, 0903.0850.

[40] T. Aaltonen et al., *First Observation of Electroweak Single Top Quark Production*, Phys. Rev. Lett. **103** (2009), 092002, 0903.0885.

[41] Tevatron Electroweak Working Group, *Combination of CDF and D0 Measurements of the Single Top Production Cross Section*, (2009), 0908.2171.

[42] `http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_public.html`.

[43] T. Aaltonen et al., *Combination of Tevatron searches for the standard model Higgs boson in the W+W- decay mode*, Phys. Rev. Lett. **104** (2010), 061802, 1001.4162.

[44] "The LEP Electroweak Working Group", 2009, `http://lepewwg.web.cern.ch/LEPEWWG/`.

[45] R. Frederix and F. Maltoni, *Top pair invariant mass distribution: a window on new physics*, JHEP **01** (2009), 047, 0712.2355.

[46] K. J. F. Gaemers and F. Hoogeveen, *Higgs Production and Decay into Heavy Flavors with the Gluon Fusion Mechanism*, Phys. Lett. **B146** (1984), 347.

[47] D. Dicus, A. Stange, and S. Willenbrock, *Higgs decay to top quarks at hadron colliders*, Phys. Lett. **B333** (1994), 126–131, hep-ph/9404359.

[48] W. Bernreuther, M. Flesch, and P. Haberl, *Signatures of Higgs bosons in the top quark decay channel at hadron colliders*, Phys. Rev. **D58** (1998), 114031, hep-ph/9709284.

[49] A. V. Manohar and M. B. Wise, *Flavor changing neutral currents, an extended scalar sector, and the Higgs production rate at the LHC*, Phys. Rev. **D74** (2006), 035009, hep-ph/0606172.

[50] M. I. Gresham and M. B. Wise, *Color Octet Scalar Production at the LHC*, Phys. Rev. **D76** (2007), 075003, 0706.0909.

[51] E. H. Simmons, *Coloron phenomenology*, Phys. Rev. **D55** (1997), 1678–1683, hep-ph/9608269.

[52] D. Choudhury, R. M. Godbole, R. K. Singh, and K. Wagh, *Top production at the Tevatron/LHC and nonstandard, strongly interacting spin one particles*, Phys. Lett. **B657** (2007), 69–76, 0705.1499.

[53] D. A. Dicus, C. D. McMullen, and S. Nandi, *Collider implications of Kaluza-Klein excitations of the gluons*, Phys. Rev. **D65** (2002), 076007, hep-ph/0012259.

[54] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, *The hierarchy problem and new dimensions at a millimeter*, Phys. Lett. **B429** (1998), 263–272, hep-ph/9803315.

[55] L. Randall and R. Sundrum, *A large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999), 3370–3373, hep-ph/9905221.

[56] L. D. Landau, Doklady Akad. Nauk, SSSR **60** (1948), 207.

[57] C.-N. Yang, *Selection Rules for the Dematerialization of a Particle Into Two Photons*, Phys. Rev. **77** (1950), no. 2, 242–245.

[58] W.-Y. Keung, I. Low, and J. Shu, *On Decays of Z' into Two Z Bosons and the Landau-Yang Theorem*, Phys. Rev. Lett. **101** (2008), 091802, 0806.2864.

[59] P. Langacker, *The Physics of Heavy Z' Gauge Bosons*, Rev. Mod. Phys. **81** (2008), 1199–1228, 0801.1345.

[60] A. Stange, W. J. Marciano, and S. Willenbrock, *Higgs bosons at the Fermilab Tevatron*, Phys. Rev. D **49** (1994), 1354–1362.

[61] C. Quigg, *Top-ology*, Phys. Today **50N5** (1997), 20–26.

[62] S. P. Martin, *A supersymmetry primer*, (1999), hep-ph/9709356.

[63] I. Aitchison, *Supersymmetry in particle physics*, Cambridge University Press, 2007.

[64] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. **D8** (1973), 1226–1239.

[65] Y. Grossman, *Phenomenology of models with more than two Higgs doublets*, Nucl. Phys. **B426** (1994), 355–384, hep-ph/9401311.

[66] A. G. Akeroyd, *Hidden top quark decays to charged Higgs scalars at the Tevatron*, (1995), hep-ph/9509203.

[67] A. Sopczak, *Limits on Higgs Boson Masses from a MSSM Parameter Scan at $\sqrt{s} \leq$202 GeV*, Tech. Report DELPHI-2001-026-MORIO-CONF-467.CERN-DELPHI-2001-026-MORIO-CONF-467, CERN, Geneva, Mar 2001.

[68] C. T. Hill and S. J. Parke, *Top production: Sensitivity to new physics*, Phys. Rev. **D49** (1994), 4454–4462, hep-ph/9312324.

[69] P. H. Frampton and S. L. Glashow, *Chiral Color: An Alternative to the Standard Model*, Phys. Lett. **B190** (1987), 157.

[70] P. H. Frampton and S. L. Glashow, *Unifiable Chiral Color with Natural GIM Mechanism*, Phys. Rev. Lett. **58** (1987), 2168.

[71] R. S. Chivukula, A. G. Cohen, and E. H. Simmons, *New strong interactions at the Tevatron?*, Phys. Lett. **B380** (1996), 92–98, hep-ph/9603311.

[72] C. T. Hill and E. H. Simmons, *Strong dynamics and electroweak symmetry breaking*, Phys. Rept. **381** (2003), 235–402, hep-ph/0203079.

[73] C. T. Hill, *Topcolor assisted technicolor*, Phys. Lett. **B345** (1995), 483–489, hep-ph/9411426.

[74] K. D. Lane, *An Introduction to technicolor*, (1993), hep-ph/9401324.

[75] M. Guchait, F. Mahmoudi, and K. Sridhar, *Tevatron constraint on the Kaluza-Klein gluon of the bulk Randall-Sundrum model*, JHEP **05** (2007), 103, hep-ph/0703060.

[76] "Visual Media Service Photo Database", `http://www-visualmedia.fnal.gov/VMS_Site/active.html`.

[77] R. J. Scherrer, *Time variation of a fundamental dimensionless constant*, (2009), 0903.5321.

[78] `http://d0server1.fnal.gov/Projects/Operations/D0RunII_DataTaking_files/`.

[79] S. Abachi et al., *The DØ detector*, (1994), FERMILAB-Pub-93/179-E.

[80] S. Abachi et al., *The DØ upgrade, the detector and its physics*, (1996), FERMILAB-Pub-96/357-E.

[81] V. M. Abazov et al., *The upgraded DØ detector*, (2006), Nucl. Instrum. Meth. A 565, 463.

[82] `http://www-d0.fnal.gov/Run2Physics/displays/presentations/`.

[83] `http://www-d0.fnal.gov/Run2Physics/WWW/drawings/d0det_tracking_ah.eps`.

[84] M. Weber, `http://d0server1.fnal.gov/projects/Silicon/www/SMT_files/talks/`.

[85] S. Abachi et al., *The D0 detector*, Nucl. Instrum. Meth. Phys. Res. **A 338** (1994), 185.

[86] V. M. Abazov et al., *The upgraded D0 detector*, Nucl. Instrum. Meth. Phys. Res. **A 565** (2006), 463, physics/0507191.

[87] D. Chapin et al., *Measurement of $z \to ee$ and $w \to e\nu$ production cross sections using one tight central electron*, (2005), DØ Note 4897.

[88] M. Voutilainen and C. Royon, *Jet p_t resolution using jetcorr v7.1*, August 2007, DØ Note 5381.

[89] `http://www-d0online.fnal.gov/www/groups/tm/tm_main.html`.

[90] G. C. Blazey, *"The DØ RunII Trigger"*, (1997), FERMILAB-CONF-97-395-E.

[91] `http://www-d0.fnal.gov/runcoor/run2best.html`.

[92] `http://www-d0.fnal.gov/d0dist/dist/packages/coor/devel/doc/coorover.ps`.

[93] A. Khanov, *HTF: histogramming method for finding tracks. The algorithm description.*, (2000), DØ Note 3778.

[94] G. Borissov, *Ordering a Chaos or...Technical Details of AA Tracking*, (2003), `http://www-d0.fnal.gov/atwork/adm/d0_private/2003-02-28/adm_talk.ps`.

[95] H. Greenlee, *The DØ Kalman Track Fit*, (2004), DØ Note 4303.

[96] A. Schwartzman and C. Tully, *Primary vertex reconstruction my means of adaptive vertex fitting*, (2005), DØ Note 4918.

[97] A. Schwartzman and M. Narain, *Vertex fitting by means of the Kalman filter technique*, (2001), DØ Note 3907.

[98] A. Garcia-Bellido et al., *Primary vertex certification in p14*, (2004), DØ Note 4320.

[99] A. Schwartzman and M. Narain, *Probabilistic primary vertex selection*, (2002), DØ Note 4042.

[100] J. Kozminski et al., *Electron likelihood in p14*, (2004), DØ Note 4449.

[101] L. Wang et al., *Electron likelihood efficiency in p17*, (2006), DØ Note 5114.

[102] P. Calfayan et al., *Muon identification certificationfor p17 data*, June 2006, DØ Note 5157.

[103] O. Brandt et al., *Muon Identification Certification for the Summer 2009 Extended Dataset (Run IIb-1 and -2)*, DØ Note 6025.

[104] `http://www-d0.fnal.gov/computing/algorithms/muon/p17/muonId_quality_type.html`.

[105] G. C. Blazey et al., *Run II Jet Physics*, (2000), hep-ex/0005012.

[106] Z. Casilum and R. Hirosky, *Jet Reconstruction Efficiency*, (1998), DØ Note 3324.

[107] G. Sterman and S. Weinberg, *Jets from Quantum Chromodynamics*, Phys. Rev. Lett. **39** (1977), 1436.

[108] W. K. Tung et al., *Heavy quark mass effects in deep inelastic scattering and global QCD analysis*, JHEP **02** (2007), 053, hep-ph/0611254.

[109] "The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups", *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006), 257, hep-ex/0509008.

[110] E. Busato and B. Andrieu, *JetAlgorithms in the DØ Run II Software: Description and User's Guide*, (2004), DØ Note 4457.

[111] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993), 187–224.

[112] S. Catani, Y. L. Dokshitzer, and B. R. Webber, *The K^- perpendicular clustering algorithm for jets in deep inelastic scattering and hadron collisions*, Phys. Lett. **B285** (1992), 291–299.

[113] S. Catani, Y. L. Dokshitzer, and B. R. Webber, *The K(T) clustering algorithm for jets in deep inelastic scattering*, Nucl. Phys. Proc. Suppl. **29A** (1992), 136–143.

[114] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993), 3160–3166, hep-ph/9305266.

[115] T. Schliephake, *Studien zu kT-Jetalgorithmen bei DØ und Vergleich mit den DØ-Conealgorithmen*, 2007.

[116] A. Harel, *Jet id optimization*, (2005), DØ Note 4919.

[117] A. Harel and R. Wagner, *Improved* L1 *confirmation*, (2005), DØ Note 4932.

[118] `http://www-d0.fnal.gov/phys_id/jes/public/plots_v7.1/`.

[119] Jet Energy Scale Working Group, *"Jet Energy Scale Determination at DØ RunII (final p17 version)"*, (2007), DØ Note 5382.

[120] D. Gilbert et al., *Measuring jet response using the missing E_{T} projection method in photon+jet events*, (2004), DØ Note 4571.

[121] N. Makovec and J.-F. Grivaz, *Shifting, smearing and removing simulated jets*, (2005), DØ Note 4914.

[122] C. Orchando and J.-F. Grivaz, *Ssr for p17*, (2008), DØ Note 5609.

[123] A. Schwartzman, *Measurement of the $b^{\pm}$ lifetime and top quark identification using secondary vertex b-tagging*, (2004), FERMILAB-THESIS-2004-21.

[124] I. Anghel, C. Gerber, E. Shabalina, and T. Ten, *Studies of taggability versus number of primary vertices for p17 data*, (2006), DØ Note 5240.

[125] A. Schwartzman and M. Narain, *Secondary vertex reconstruction using the kalman filter*, (2001), DØ Note 3908.

[126] D. Boline, L. Feligioni, and M. Narain, *Update on b-quark jet identification with secondary vertex reconstruction using DØ reco version p14-pass2 samples*, (2005), DØ Note 4796.

[127] C. Gerber, E. Shabaline, and G. O. y Garcon, *Taggability in p14 pass2 data*, (2006), DØ Note 4995.

[128] Y. Peters, *Measurements and searches with top quarks - Measurement of the top quark pair production cross section, the ratio of branching fractions and searches for new physics in the top quark sector with the DØ experiment*, Ph.D. thesis, Wuppertal, June 2008.

[129] K. Hanagaki and J. Kasper, *Identification of b-jet by soft muon*, (2005), DØ Note 4867.

[130] R. Demina et al., *Measurement of b-tagging efficiency and mis-tagging rates with CSIP method*, (2004), DØ Note 4432.

[131] R. Demina, A. Khanov, and F. Rizatdinova, *b-tagging with counting signed impact parameter method*, (2002), DØ Note 4049.

[132] D. Bloch and B. Clément, *Update of the JLIP b-tagger performance in p14/pass2 with JES 5.3*, (2005), DØ Note 4824.

[133] B. Clément, D. Bloch, D. Gelé, S. Greder, and I. Ripp-Baudot, *Performance of the JLIP b-tagger in p14*, (2004), DØ Note 4348.

[134] A. Schwartzman and M. Narain, *b quark jet identification via secondary vertex reconstruction*, (2003), DØ Note 4080.

[135] Y. Peters, H. Greenlee, A. Haas, and A. Schwartzman, *Optimization of the secondary vertex tagger in p17*, (2006), DØ Note 5265.

[136] A. Schwartzman and M. Narain, *b quark jet identification via secondary vertex reconstruction in DØ reco p13 software*, (2003), DØ Note 4081.

[137] A. Schwartzman and M. Narain, *Secondary vertex reconstruction using the Kalman filter*, (2001), DØ Note 3908.

[138] T. Scanlon, *A neural network b-tagging tool*, (2005), DØ Note 4889.

[139] M. Anastasoaie, S. Robinson, and T. Scanlon, *Performance of the NN b-tagging Tool on p17 Data*, 2006, DØ Note 5213.

[140] T. Gadfort and others, *Performance of the dzero nn b-tagging tool on p20 data*, (2007), DØ Note 5554.

[141] C. DooKee et al., *Measurement of the $t\bar{t}$ Production Cross-Section at $\sqrt{s} = 1.96$ TeV in the Lepton+Jets Final State using Neural Network b-tagging algorithm on $1\,fb^{-1}$ of DØData*, (2007), DØ Note 5335.

[142] B. Clement et al., *SystemD or how to get signal, backgrounds and their efficiencies with real data*, (2001), DØ Note 4159.

[143] S. Calvet, P. Verdier, and E. Kajfasz, *Towards $\not\!\!E_{\mathrm{T}}$ certification and unclustered energy studies*, (2005), DØ Note 4927.

[144] G. Bernardi, E. Busato, and J.-R. Vlimant, *Improvements from the T42 Algorithm on Calorimeter Object Reconstruction*, (2004), DØ Note 4335.

[145] C. DooKee et al., *Measurement of the $t\bar{t}$ production cross-section at $\sqrt{s} = 1.96$ TeV in the lepton+jets final state using neural network btagging algorithm on $3.7\,fb^{-1}$ of DØData*, (2009), DØ Note 5528.

[146] M. Voutilainen, *Single jet trigger efficiencies in Run IIa*, December 2007, DØ Note 5549.

[147] M. Hohlfeld, *Measurement of Single Electron Trigger Efficiencies for RunIIB*, October 2008, DØ Note 5783.

[148] F. Deliot and V. Sharyy, *Trigger Efficiency Measurement for the OR of Electron Muon Triggers*, February 2008, DØ Note 5587.

[149] Application Software Group (CERN), `CERN` Program Library Long Writeup W5013, `http://wwwasd.web.cern.ch/wwwasd/geant/`.

[150] S. Agostinelli et al., *Geant4: A simulation toolkit*, (2003), Nucl. Instrum. Meth. **A 506**, 250.

[151] Y. Fisyak and J. Womersley, *D0 geant simulation of the total apparatus response*, (1997), DØ Note 3191.

[152] DØ Collaboration, *D0sim user manual*, (1986), 407.

[153] `http://www-d0.fnal.gov/computing/MonteCarlo/MonteCarlo.html`.

[154] J. Pumplin et al., *New generation of parton distributions with uncertainties from global qcd analysis*, JHEP **07** (2002), 012, hep-ph/0201195.

[155] Y. Peters et al., *Simultaneous measurement of $B(t \to Wb)/B(t \to Wq)$ and $\sigma(p\bar{p} \to t\bar{t}) * b(t \to wq)^2$ with p17 data*, (2007), DØ Note 5422.

[156] D. Wicke and M. Vaupel, *Search for $t\bar{t}$ resonance in the lepton+jets final state in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,TeV$*, 2007, DØ Note 5434.

[157] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003), 001, hep-ph/0206293.

[158] S. Höche et al, *Matching parton showers and matrix elements*, (2006), hep-ph/0602031.

[159] E. E. Boos et al., *Method for simulating electroweak top-quark production events in the NLO approximation: SingleTop event generator*, Phys. Atom. Nucl. **69** (2006), 1317.

[160] J. Hays et al., *Single electron efficiencies in p17 data and monte carlo using p18.05.00 d0correct*, (2006), DØ Note 5105.

[161] S. Park and M. Begel, *Efficiency of the data quality calorimeter flags*, (2007), DØ Note 5324.

[162] H. Schellman, *The longitudinal shape of the luminosity region*, (2006), DØ Note 5142.

[163] `http://www-d0.fnal.gov/Run2Physics/top/d0_private/wg/triggers/triggers.html`.

[164] Y. Peters, M. Begel, K. Hamacher, and D. Wicke, *Reweighting of the fragmentation function for the DØ monte carlo*, 2007, DØ Note 5325.

[165] B. Martin, *Z p_T reweighting procedure for p17 Alpgen Monte Carlo*, (2008), DØ Note 5571.

[166] D. Bandurin et al., *Electron and photon efficiencies for the re-processed p17 and p20 data*, (2008), DØ Note 5798.

[167] M.-A. Pleier, *Measurement of the electron and muon fake rates in lepton+jets datasets*, (2007), DØ Note 5469.

[168] E. Barberis et al., *The matrix method and its error calculation*, (2004), DØ Note 4564.

[169] B. Martin et al., *Measurement of the $t\bar{t}$ production cross section at $\sqrt(s) =$ 1.96 TeV in the ee final state using p17 data set*, (2008), DØ Note 5579.

[170] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl, *The Fully differential single top quark cross-section in next to leading order QCD*, Phys. Rev. **D66** (2002), 054024, hep-ph/0207055.

[171] A. J. Tanasijczuk, *Search for Single Top Quark Production at the D0 Experiment using Bayesian Neural Networks*, (2009), 0906.0119.

[172] Y. Arnoud et al., *Measurement of the ttbar Production Cross-section at $\sqrt(s) = 1.96$ TeV in the Dilepton Final States Using 1 fb^{-1}*, DØ Note 5371.

[173] S. Fu and A. Haas, *A search for $ZH(\rightarrow e^+e^-b\bar{b})$ Production with the DØ Detector in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, (2006), DØ Note 5275.

[174] S. Fu, *A Search for $ZH(\rightarrow e^+e^-b\bar{b})$ Production with the DØ Detector in $p\bar{p}$ Collisions at $\sqrt{s} =$1.96 TeV*, DØ Note, no DØ Note number yet .

[175] G. J. O. y Garzon et al., *Measurement of the $t\bar{t}$ production cross-section at $\sqrt{s} = 1.96$ TeV measurement of the ttbar production cross section in pp-bar collisions at $\sqrt(s) = 1.96$ TeV using secondary vertex b-tagging*, (2006), DØ Note 5113.

[176] "Top Physics Working Group", *Direct measurement of the top quark mass in the lepton+jets channel using Run II data*, (2004), DØ Note 4426.

[177] V.M. Abazov et al., *Direct Measurement of the Top Quark Mass*, Phys. Rev. Lett. **79** (1997), 1197.

[178] J. P. Berge, F. T. Solmitz, and H. D. Taft, Rev. of Scientific Instruments **32 (5)** (1961), 538–548.

[179] S. S. Snyder, *Measurements of the top-quarks mass at DØ*, Ph.D. thesis, SUNY, Stony Brook, May 1995.

[180] M. Vaupel, *Measurement of the $t\bar{t}$ invariant Mass Distribution and Search for $t\bar{t}$ Resonances*, Ph.D. thesis, Wuppertal, March 2006.

[181] A. Harel and D. Gilbert, *How to vary jet id efficiencies, energy scale and resolution*, (2008), `https://plone4.fnal.gov/P1/D0Wiki/physics/top/HowToVaryJetCorrections/view?searchterm=JSSR`.

[182] Tevatron Electroweak Working Group, *Combination of CDF and D0 Results on the Mass of the Top Quark*, (2008), arXiv:0808.1089 [hep-ex].

[183] T. L. Edwards et al., *Determination of the effective inelastic $p\bar{p}$ cross-section for the D0 Run II luminosity measurement*, FERMILAB-TM-2278-E.

[184] private communication with Y. Peters.

[185] M. Arthaud et al., *Measurement of the ttbar production cross-section at sqrts=1.96tev in electron muon final states using p17 data set*, (2008), DØ Note 5580.

[186] M. Agelou et al., *Search for single top quark production at DØ in RunII*, (2004), DØ Note 4398.

[187] M. Agelou et al., *Improved search for single top quark production*, (2005), DØ Note 4670.

[188] I. Bertram et al., *A recipe for the construction of confidence limits*, FERMILAB-TM-2104.

[189] S. Jain and H. B. Prosper and R. Schwienhorst, *Statistical methods implemented in the package "top_ statistics"*, October 2009, DØ Note 5817.

[190] T. Aaltonen et al., *Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV*, (2008).

[191] T. Aaltonen et al., *Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV*, Phys. Rev. Lett. **100** (2008), 231801, 0709.0705.

[192] V. M. Abazov et al., *Precise measurement of the top quark mass from lepton+jets events at D0*, Phys. Rev. Lett. **101** (2008), 182001, 0807.2141.

Acknowledgements

First I have to thank Prof. Mättig for giving the possibility and subject of the thesis. I wish to thank Prof. Wolfgang Wagner for taking the burden in reading the thesis as a co-corector and answering many physics question related and unrelated to the thesis.

Of course, my supervisor Daniel Wicke needs to be thanked for the energy and time he spared on my thesis. Though having a rough start, I very much appreciated the work done together and would like to thank him for the help and encouragement I received during countless discussions on the thesis, physics in general and limit setting procedures in particular.

I would like to thank Klaus Hamacher for increasing my physic knowledge by giving his own point of view about ideas on how a subject should be approached. The increase of my physical health by swimming should not pass unmentioned either. I also wish to express that Klaus Hamacher is an important part of this working group and succeeds in turning it into a good place to work.

I would like to thank my office colleague Yvonne Peters for many discussions about almost everything and bearing to share the same office in Wuppertal and Fermilab. Apart from that, our co-operation has been very successful in developing a more convenient way of data processing.

I wish to thank Hendrik for his advice on root and linux in general, Marcello for his Python skripts and for being a good office colleague. Markus for sharing his theoretical insight in maths and physics as well as private time. Prof. Zeitnitz for physics and teaching helps and the chance to get back electronics once more. Malgorzata Worek for giving a real good theoretical insight in the Standard Model and Neutrino mixing/oscillation. More colleagues in Wuppertal helped making the stay here enjoyable and I wish to thank them all.

Besides Wuppertal, the Fermilab colleagues need to be thanked, too. I wish to highlight Alan Jonckheere for making it possible to overcome the administrative barrier to make computing work and countless questions about the infrastructure.

Also my roommate, Edgar Carrera, with whom I spent a year at Naperville, needs to be thanked for enlighting my time in the USA and caring enough to visit me in the hospital there.

I wish to thank my girlfriend Melanie Soodt for enduring my temper and having patience especially during the last part of the thesis and, of course, for supporting me. I wish to thank Anna Hennig for sharing my life for more than seven years and her support during the first part of the thesis. Especially my parents and grandmother need to be thanked who have always supported me and relinquished on spending time with me in favor of this thesis.