Jet energy calibration and a search for supersymmetry with vector boson fusion channel like sign di-τ_h final states

Dissertation

zur Erlangung des Doktorgrades

des Departement Physik

der Universität Hamburg

vorgelegt von

Denis Rathjens

Hamburg, 2015

The more accurate the map, the more it resembles the territory. The most accurate map possible would be the territory, and thus would be perfectly accurate and perfectly useless.

From the notebooks of Mr. Ibis out of the book "American Gods" by Neil Gaiman.

UNIVERSITY OF HAMBURG

Abstract

Fakultät für Mathematik, Informatik & Naturwissenschaften
Fachbereich Physik

by Denis Rathjens

At the LHC, the production of jets has the highest cross section out of all processes. Therefore, jets are important objects for calibration, reconstruction and identification at the CMS experiment.

In this thesis, the calibration of the jet energy scale with respect to residual differences between data and simulation after simulation-based precalibrations is shown. A correction for the $\sqrt{s} = 8$ TeV run of 2012 depending on jet transverse momentum and pseudorapidity is derived using di-jet final states.

Furthermore, the capacity of jets to be misidentified as hadronically decaying τ leptons is demonstrated. A method for an approximate simulation based description of this property is shown in the context of a search for supersymmetry in vector boson fusion final states.

Am LHC haben Ereignisse mit jets den höchsten Produktionswirkungsquerschnitt. Daher sind jets eines der wichtigsten Objekte zur Kalibration, Rekonstruktion und Identifikation am CMS-Experiment.

In dieser Arbeit wird die Kalibration der Jetenergieskala hinsichtlich der verbliebenen Unterschiede zwischen Daten und Simulation nach aus der Simulation bestimmten Korrekturen gezeigt. Eine Korrektur, bestimmt aus Zweijetereignissen für die Datennahme von 2012 bei $\sqrt{s} = 8$ TeV, die vom transversen Impuls und der Pseudorapidität der Jets abhangt, wird vorgestellt.

Weiterhin wird die Eigenschaft von jets dargestellt, als hadronisch zerfallende τ-Leptonen fehlrekonstruiert zu werden. Eine Methode einer näherungsweisen simulationsbasierten Beschreibung dieser Eigenschaft wird im Kontext einer Suche nach Supersymmetrie mit Endzuständen, bei denen es zu einer Fusion von Vektorbosonen kommt, vorgestellt.

Contents

Chapter 1

Introduction

Since mankind was capable of dividing the world into "self" and "everything else", it has struggled with defining what "everything else" is supposed to be. Many ways to tackle this problem have been invented and random correlations between occurences, hearsay and/or imagination have been mistaken for causation and become "common knowledge" for millenia. The drive of man to find a "sense" in how things are, and therefore to believe in what makes "sense", is as limitless as its imagination. Today, the most successful system of interpreting the world other than ourselves is the scientific method, an inherent disbelief system. As formulated in Ref. [2], the basic idea is to disbelieve everything, until mounting reproducible evidence overpowers disbelief.

This method, so very much against human nature, is what brought the scientific advances of the last half millenium. In line with this philosophy, the belief in the most successful theory of elementary particles and their interactions, the Standard Model of particle physics (SM) (see Sec. 2.1), needs to be tried and tested wherever and however possible.

The Large Hadron Collider (LHC) (see Sec. 3.1) was built to extend the knowledge about the SM into new energy regimes and find its missing predicted constituent, the Higgs boson. As a proton-proton collider, the prevalent processes are Quantum Chromodynamics (QCD) processes that constitute the least understood and theoretically most complicated part of the SM. Since quarks and gluons hadronize almost immediately after production, the observable phenomena in a detector are jets (see Sec. 4.2) that are subject to various detector effects and theoretical uncertainties. In this thesis, part of the calibration chain to relate jets as experimental objects to simulated objects that are comparable to theory is undertaken. As such, almost all analyses within the Compact Muon Solenoid (CMS) (see Sec. 3.2) collaboration working with jet objects and $\sqrt{s} = 8$ TeV data utilize the part of the jet energy corrections presented in this thesis. It is of paramount importance to get as diligent and exact a result on jet calibration as possible,

because the uncertainties of the jet energy measurement propagate to the uncertainty of all analyses using jets. To this effect, the $\sqrt{s} = 8$ TeV jet energy corrections (JEC) manage to use improved methods and superior event numbers in comparison to the $\sqrt{s} = 7$ TeV variants to facilitate the most precise JEC determination to date.

Such a very precise calibration allows equally precise measurements of SM properties. Despite no clear counterindication in precision measurements of the phasespace described by the SM, the SM itself is far from being a complete theory. Gravity is not included, neutrinos have mass, the renormalization of scalar masses is quadratically divergent, and neither dark energy nor dark matter hinted at by astronomical measurements are included. Therefore, extensions to the SM like supersymmetry (SUSY) (see Sec. 2.2) that could alleviate at least some of the aforementioned problems are theoretically formulated and need to be excluded, unless they exist. The second part of this thesis deals with vector boson fusion (VBF) electroweak production of compressed SUSY spectra. Specifically, decay chains with charginos or heavy neutralinos decaying to $\bar{\nu}_\tau \tilde{\tau} \rightarrow \tau \tilde{\chi}_1^0$ at 100% branching ratios are examined (see Sec. 6.1). Tagging VBF jets allows a unique access to previously inaccessible parts of SUSY phasespace, but also provides new challenges.

In this thesis, the like sign di-hadronic τ channel and especially the validation of its background determination methodology in simulation is introduced (see Sec. 6.2). The main background contribution with very high purity in the signal region is four-jet QCD. As NNLO calculations have not been available at the time the work has been conducted, large theoretical uncertainties on any estimation from simulation would have been needed to be imposed. Instead, the normalization of the background contribution is estimated from data, directly. While this allows for a reduction of theoretical uncertainties, it disallows the estimation of underlying assumptions. Such assumptions are made, e.g., on the correlation or lack thereof of the vector boson fusion jet system observables to hadronic τ observables. In order to estimate these contributions on an inclusive multijet sample, one cannot directly take the simulation at face value, as the stringent selection requirements of two jets being misidentified as hadronically decaying τ plus two additional jets at large rapidities is not passed by a single event in the aforementioned samples. Instead, a method is developed to estimate the probability of jets to be misreconstructed as hadronic τ in simulation. Using this method, simulated samples are reweighted and reinterpreted as if they would pass the requirement of two hadronic τ, in order to estimate systematic uncertainties on the background estimation in data. Optimising the analysis based on the systematic uncertainty evaluation, a relative total uncertainy of $\approx 20\%$ on the background prediction that is competitive with less challenging τ decay channels has been achieved. Finally, the analysis contributed to a 95% confidence level (CL) exclusion of $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} < 170$ GeV in a compressed mass spectra scenario of $m_{\tilde{\chi}_1^\pm} - m_{\tilde{\chi}_1^0} = 50$ GeV and $\Delta m(\tilde{\chi}_1^\pm, \tilde{\tau}) = 5$ GeV, a formerly

uncharted region of phasespace.

This thesis is organized by introducing the SM, motivating SUSY and showing the first VBF search results of all channels combined in Chap. 2. Then, the LHC accelerator and the CMS detector are introduced in Chap. 3. All reconstructed physical objects used in this thesis are detailed in Chap. 4. The $\sqrt{s} = 8$ TeV JEC are introduced in Chap. 5 and the relative residual correction with dijets is determined. Finally, Chap. 6 introduces the VBF like sign di-hadronic τ channel and demonstrates the determination of systematic uncertainties with dicing techniques in simulated events. Chap. 7 summarizes and concludes this thesis.

Chapter 2

Theoretical foundations

Two fundamental theories describe the four fundamental forces known to mankind. The electromagnetic, weak and strong forces are described by the Standard Model (SM) of particle physics summarized in the following section. On the other hand, the most obvious force in everyday life, gravity, is described by the theory of general relativity and too weak on a particle level to be measured by particle physics at the present time. Therefore, it will be omitted, here.

Still, there are other shortcomings of the SM that require the introduction of new physics theories, like dark matter or quadratic higgs mass renormalization divergencies. In order to fix such problems, supersymmetry (SUSY) as an extension of the SM will be introduced with an emphasis on vector boson fusion (VBF) channel production and the theoretical motivation for searching for this kind of SUSY.

2.1 The Standard Model of particle physics

As a theoretical construct, the SM is a renormalizable, locally gauge invariant theory. In the standard model, indivisible indistinguishable elementary particles are formulated in the vocabulary of quantum field theory.

There are two general classes of particles, characterized by their spin. Half-integer spin particles are fermions, integer spin particles are bosons. Fermions are subject to the Pauli exclusion principle disallowing two fermions from entering the same state in the same space at the same time. A full overview of all particles contained in the SM is shown in Fig. 2.1.
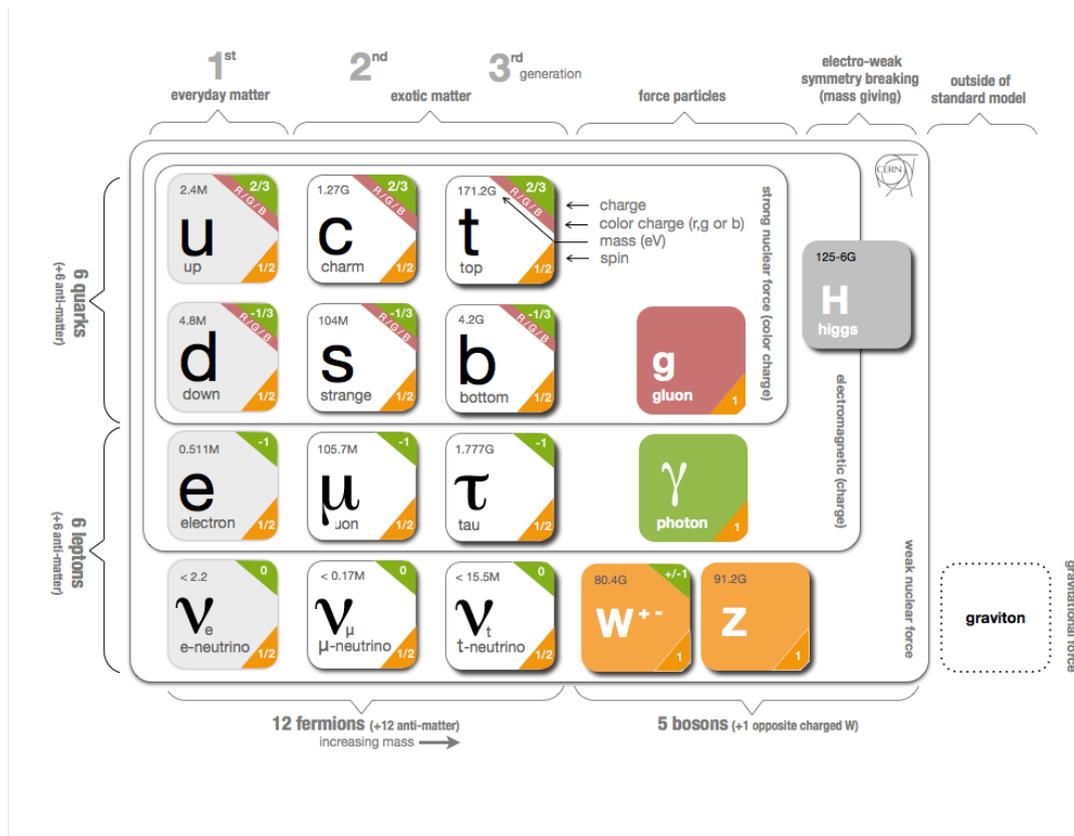
FIGURE 2.1: Taken from Ref. [3], the figure shows the particles, their proper-
ties and associated forces described in the SM, as well as the theoretical force
exchange particle for gravity, outside of the SM.

As explained in, e. g. , Ref. [4] that is taken as a guideline for this section, information of
the SM is mathematically formulated as a Lagrangian density ($\mathcal{L}$) which is the quantum
field equivalent of a classical langrangian function $L = T - U$ with a kinetic energy T
and a scalar potential energy U. Both, the classical Lagrangian function and the field
theory Lagrangian density allow for the derivation of equations of motion in generalized
coordinates by applying the Euler-Lagrange equation. In the classical case with general-
ized coordinates q_i and their time derivatives $\dot{q}_i$, the formulation of the Euler-Lagrange
equation is shown in Eq. 2.1.

$$\frac{\mathrm{d}}{\mathrm{dt}} \left(\frac{\delta \mathrm{L}}{\delta \dot{q}_i} \right) = \frac{\delta \mathrm{L}}{\delta q_i} \tag{2.1}$$

A Lagrangian density is a function not of generalized coordinates, but of generalized
fields $\phi_i(x, y, z, t)$ with $(x, y, z, t) = x^\mu$ introduced as a notational shorthand. The re-
spective relativistic formulation of the Euler Lagrange equation is shown in Eq. 2.2.

$$\delta_\mu \left(\frac{\delta \mathcal{L}}{\delta(\delta_\mu \phi_i)} \right) = \frac{\delta \mathcal{L}}{\delta \phi_i} \tag{2.2}$$

This way, equations of motion for particles can be generated by knowing the appropriate Lagrangian density and this density contains all the physical information about the quantized fields involved. The different forces known in the framework of the SM are the electromagnetic, the weak and the strong force and will be introduced with their Lagrangian densities in the following chapters. Gravity is not included in the SM.

2.1.1 Quantum Electrodynamics

While in classical mechanics a lagrangian function L can be derived, the most simple part of the SM is chosen such as to reproduce classical field theory in a classical limit for the best-known interaction, the electromagnetic force. Using spinor fields ψ with their adjoints $\overline{\psi}$, the Dirac Lagrangian out of Eq. 2.3 is used as a starting point.

$$\mathcal{L} = i\hbar c \overline{\psi} \gamma^\mu \delta_\mu \psi - mc^2 \overline{\psi}\psi \tag{2.3}$$

It is invariant under global phase transformations $\psi \to e^{i\phi}\psi$ with $\phi \in \mathcal{R}$. One can now require that local phase transformations $\psi \to e^{i\phi(x)}\psi$ should hold as well. This is not true out of the box, because $\mathcal{L} \to \mathcal{L} - \hbar c(\delta_\mu \phi)\overline{\psi}\gamma^\mu\psi$ is the result of the local transformation. One can set $\phi(x) = -\frac{q}{\hbar c}\lambda(x)$ and define q to be a particle charge. Still, the Lagrangian density of quantum electrodynamics (QED) needs to be modified, in order to make it invariant under local phase transformation. So besides the kinematic term $i\hbar c \overline{\psi}\gamma^\mu \delta_\mu \psi$ and the potential term $mc^2\overline{\psi}\psi$, a field term is to be introduced. Defining a vector field A_μ that follows the transformation rule $A_\mu \to A_\mu + \delta_\mu \lambda$, this extra derivative would cancel the extra derivative from $\mathcal{L} \to \mathcal{L} + (q\overline{\psi}\gamma^\mu\psi)\delta_{mu}\lambda$. Involving a new vector field also involves new terms for the propagation of that field including a Proca Lagrangian as shown in Eq. 2.4 with the shorthand $F^{\mu\nu} = (\delta^\mu A^\nu - \delta^\nu A^\mu)$.

$$\mathcal{L} = \frac{-1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}\left(\frac{m_A c}{\hbar}\right)^2 A^\nu A_\nu \tag{2.4}$$

Since $A^\nu A_\nu$ is not invariant under local phase transformation, the vector field needs to be massless, in order to fulfill local phase transformation invariance.

Identifying A^μ as the electromagnetic potential, the full Lagrangian density for quantum electrodynamics is shown in Eq. 2.5.

$$\mathcal{L} = [i\hbar c \overline{\psi}\gamma^\mu \delta_\mu \psi - mc^2\overline{\psi}\psi] - \left[\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\right] - (q\overline{\psi}\gamma^\mu\psi)A_\mu \tag{2.5}$$

It describes the interaction of particles carrying electromagnetic charge q via massless photons.

This part can be interpreted as a transformation by multiplication with a unitary 1×1

matrix $U = e^{i\phi}$ like $\psi \to U\psi$ with $U^\dagger U = 1$. Therefore, QED is classified as $U(1)$ gauge invariant.

Alternatively, one can redefine the derivative by a "covariant derivative" $\mathcal{D}_\mu = \delta_\mu + \frac{iq}{\hbar c} A_\mu$, so that Eq. 2.6 is a shorthand full description of QED.

$$\mathcal{L} = i\hbar c \overline{\psi} \gamma^\mu \mathcal{D}_\mu \psi - mc^2 \overline{\psi}\psi - \left[\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} \right] \tag{2.6}$$

2.1.2 Quantum Flavordynamics

The weak force is described by Quantum Flavordynamics (QFD). This theory is classified as invariant under $SU(2)$ transformations $\psi \to U\psi$ with a unitary 2×2 matrix U with determinant 1 where wave functions are defined as two component spinors $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ and together with a mass matrix $M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}$ form the Lagrangian in Eq. 2.7.

$$\mathcal{L} = i\hbar c \overline{\psi} \gamma^\mu \delta_\mu \psi - c^2 \overline{\psi} M \psi \tag{2.7}$$

Introducing the Pauli matrices τ_i as specified in Eq. 2.8 and constants $a_{1..3}$, a global $SU(2)$ transformation can be written as $\psi \to e^{i\tau \cdot a}\psi$ and summarily a local $SU(2)$ transformation is written as $\psi \to e^{-iq\tau \cdot \lambda(x)/\hbar c}\psi$.

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.8}$$

Applying that to the Lagrangian shows that it is, again, not invariant. As for QED, a covariant derivative $\mathcal{D}_\mu = \delta_\mu + \frac{iq}{\hbar c} \tau \cdot \mathbf{A}_\mu$ can be defined with three new vector fields $\mathbf{A}_\mu$. Analogous to the QED case, the vector fields will require their own Lagrangian, as shown in Eq. 2.9.

$$\mathcal{L}_A = -\frac{1}{16\pi} \mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} + \frac{1}{8\pi} \left(\frac{m_A c}{\hbar} \right)^2 \mathbf{A}^\nu \cdot \mathbf{A}_\nu \tag{2.9}$$

Again, the mass term is not locally gauge invariant and so, given no further modification, force carrying bosons would need to be massless. In addition, the $\mathbf{F}^{\mu\nu} = \delta^\mu A^\nu - \delta^\nu A^\mu$ relation is not invariant in this case. Instead, the relation of Eq. 2.10 is taken.

$$\mathbf{F}^{\mu\nu} = \delta^\mu \mathbf{A}^\nu - \delta^\nu \mathbf{A}^\mu - \frac{2q}{\hbar c} (\mathbf{A}^\mu \times \mathbf{A}^\nu) \tag{2.10}$$

With all these modifications, the resulting Lagrangian density is shown in Eq. 2.11.

$$\mathcal{L} = i\hbar c \overline{\psi} \gamma^\mu \mathcal{D}_\mu \psi - c^2 \overline{\psi} M \psi - \frac{1}{16\pi} \mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} \tag{2.11}$$

It has the special property that it allows flavor changing charged currents. For this effect, parity has to be maximally violated, meaning the introduction of an axial vector coupling. The $W^{\pm}$ bosons have axial vector couplings as formulated in Eq. 2.12 denoting a vertex factor.

$$\frac{-ig_W}{2\sqrt{2}}\gamma^{\mu}(1-\gamma^5) \tag{2.12}$$

For the Z^0 boson, a vector-like coupling c_V^f and an axial-vector-like coupling c_A^f exist, as formulated in Eq. 2.13 with the respective couplings for different fermions f being listed in Tab. 2.1.

$$\frac{-ig_Z}{2}\gamma^{\mu}(c_V^f - c_A^f\gamma^5) \tag{2.13}$$

f	c_V^f	c_A^f
$\nu_e,\ \nu_\mu,\ \nu_\tau$	$\frac{1}{2}$	$\frac{1}{2}$
$e,\ \mu,\ \tau$	$-\frac{1}{2} + 2\sin^2(\theta_W)$	$-\frac{1}{2}$
$u,\ c,\ t$	$\frac{1}{2} - \frac{4}{3}\sin^2(\theta_W)$	$\frac{1}{2}$
$d,\ s,\ b$	$-\frac{1}{2} + \frac{2}{3}\sin^2(\theta_W)$	$-\frac{1}{2}$

TABLE 2.1: Vector couplings c_V^f and axial-vector couplings c_A^f of the Z^0 boson to fermions.

While the formulation is sound, so far, QFD in nature is a broken symmetry, because the three force-carrying bosons associated with it, the $W^{\pm}$ and Z^0, are all massive and would therefore not fit into the theory without modifications. This has been resolved by the Higgs mechanism (see Ref. [5]) as discussed in the following subsection.

2.1.3 Electroweak symmetry breaking: The Higgs Mechanism

So far, the ground state energy of all the theories formulated in the previous subsections has been considered to be equal to zero. The general idea behind the Higgs mechanism is to dynamically generate a mass term for vector fields that is gauge invariant by introducing a potential with a nonzero ground state. Therefore, a two component potential $\phi = \phi_1 + i\phi_2$ (called an isospin doublet) is introduced and the globally gauge invariant starting point is shown in Eq. 2.14.

$$\mathcal{L} = \frac{1}{2}(\delta_\mu\phi)^*(\delta^\mu\phi) + \frac{1}{2}(\phi^*\phi) - \frac{1}{4}(\phi^*\phi)^2 \tag{2.14}$$

As before, a covariant derivative $\mathcal{D}_\mu = \delta_\mu + \frac{iq}{\hbar c} A_\mu$ is defined and a field propagator added to the Lagrangian, as shown in Eq. 2.15.

$$\mathcal{L} = \frac{1}{2}\left[\left(\delta_\mu - \frac{iq}{\hbar c}A_\mu\right)\phi^*\right]\left[\left(\delta^\mu + \frac{iq}{\hbar c}A^\mu\right)\phi\right] + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} \tag{2.15}$$

Defining new fields $\eta = \phi_1 - \mu/\lambda$ and $\chi = \phi_2$, a very complicated expression shown in Eq. 2.16, 2.17, 2.18, 2.18 and 2.20 results.

$$\mathcal{L} = \left[\frac{1}{2}(\delta_\mu\eta)(\delta^\mu\eta) - \mu^2\eta^2\right] + \left[\frac{1}{2}(\delta_\mu\chi)(\delta^\mu\chi)\right] \tag{2.16}$$

$$+ \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \left(\frac{q\mu}{\sqrt{2}\hbar c\lambda}\right)^2 A_\mu A^\mu\right] \tag{2.17}$$

$$+ \left\{\frac{q}{\hbar c}[\eta(\delta_\mu\chi) - \chi(\delta_\mu\eta)]A^\mu + \frac{\mu}{\lambda}\left(\frac{q}{\hbar c}\right)^2 \eta(A_\mu A^\mu)\right. \tag{2.18}$$

$$\left. + \left(\frac{q}{\sqrt{2}\hbar c}\right)^2 (\chi^2 + \eta^2)(A_\mu A^\mu) - \lambda\mu(\eta^3 + \eta\chi^2) - \frac{1}{4}\lambda^2(\eta^4 + 2\eta^2\chi^2 + \chi^4)\right\} \tag{2.19}$$

$$+ \left(\frac{\mu q}{\lambda\hbar c}\right)(\delta_\mu\chi)A^\mu + \left(\frac{\mu^2}{2\lambda}\right)^2 \tag{2.20}$$

The different terms have different consequences. Eq. 2.16 represents a scalar particle η of mass $\sqrt{2}\mu\hbar/c$ that may identify as the recently discovered new boson (see Ref. [6]), as well as a massless Goldstone boson χ. Eq. 2.17 contains the gauge field A^μ that has now a mass term $m_A = \frac{2\sqrt{\pi}q\mu}{\lambda c^2}$. The Eq. 2.18 and 2.19 specify the different couplings to χ, η and A^μ.

Finally, choosing a different gauge $\phi \to \phi' = (\cos\theta + i\sin\theta)(\theta_1 + i\theta_2)$ with $\theta = -\tan^{-1}(\phi_2/\phi_1)$, the Goldstein boson χ can be removed, completely. But this does not change the content of the Lagrangian.

2.1.4 Electroweak unification

Now that there is a principle way to dynamically generate massive vector fields and therefore have massive gauge bosons, one more ingredient is to unify QED and QFD in a $U(1) \times SU(2)$ gauge theory. According to Ref. [7], one can define new gauge fields $W_{1..3\mu}$ and B_μ that mix just in such a way that a massless photon A_μ in Eq. 2.21, and massive charged W-bosons $W_\mu^\pm$ in Eq. 2.22, as well as massive charged neutral Z-boson Z_μ^0 in Eq. 2.23 result.

$$A_\mu = \sin\theta_W W_{3\mu} + \cos\theta_W B_\mu \tag{2.21}$$

$$W_\mu^\pm = (W_{1\mu} \mp iW_{2\mu})/\sqrt{2} \tag{2.22}$$

$$Z_\mu^0 = -\cos\theta_W W_{3\mu} + sin\theta_W B_\mu \tag{2.23}$$

Defining three weak isospin currents $\mathbf{j}_\mu$ (see Eq. 2.24), a weak hypercharge current j_μ^Y (see Eq. 2.25), and the hypercharge Y (see Eq. 2.26), one can assemble the electroweak gauge boson sector with the isotriplet $\mathbf{W}^\mu$ identical to the fields A^μ out of Sec. 2.1.2 and an isosinglet B^μ, as shown in Eq. 2.27.

$$\mathbf{j}_\mu = \frac{1}{2}\overline{\chi}_L \gamma_\mu \boldsymbol{\tau} \chi_L \tag{2.24}$$

The expression χ_L denotes any left-handed weak isospin doublets, e.g., electron and electron-neutrino or up and down quark.

$$j_\mu^Y = 2j_\mu^{em} - 2j_\mu^3 \tag{2.25}$$

Here, j_μ^3 is the neutral weak current and j^{em} is the electric current.

$$Y = 2(Q - I^3) \tag{2.26}$$

For the hypercharge, Q is the electric charge and I^3 the third component of the isospin.

$$-i\left[g_W \mathbf{j}_\mu \cdot \mathbf{W}^\mu + \frac{g'}{2}j_\mu^Y B^\mu\right] \tag{2.27}$$

The two charged states represent the $W^\pm$ boson. Because the two neutral states W^3 and B mix, the photon as a massless linear combination and the orthogonal massive combination, the Z^0 boson, result, as indicated in the beginning. These four particles constitute, together with the Higgs boson, the electroweak sector.

The angle θ_W describes the Weinberg mixing angle. The electromagnetic coupling is $q = g\sin\theta_W$p, the weak hypercharge coupling is $g' = g\tan\theta_W$ and one can also define the relations $\sin\theta_W = g'/\sqrt{g^2 + g'^2}$ and $\cos\theta_W = g/\sqrt{g^2 + g'^2}$

2.1.5 Quantum Chromodynamics

The strong force is described by Quantum Chromodynamics (QCD). This theory is classified as invariant under $SU(3)$ transformations $\psi \to U\psi$ with a unitary 3×3 matrix U with determinant 1. For QCD, three component spinors $\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_c \end{pmatrix}$ with red, green and blue color charge are defined. Analogous to the previous gauge theories, the starting Lagrangian density is shown in Eq. 2.28.

$$\mathcal{L} = i\hbar c\overline{\psi}\gamma^\mu \delta_\mu \psi - mc^2\overline{\psi}\psi \tag{2.28}$$

As $U(1)$ has one degree of freedom and no condition, there is one vector field. For $SU(2)$ four degrees of freedom and one special condition, there are $2^2 - 1 = 3$ vector fields, the $W^\pm$ and Z^0. Accordingly, a $SU(3)$ theory needs to introduce eight generators $\boldsymbol{\lambda} = \lambda_{1..8}$, namely the Gell-Mann matrices shown in Eq. 2.29, 2.30 and 2.31, and can allow nine real numbers $\mathbf{a} = a_{1..8}$ and θ.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.29}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{2.30}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{2.31}$$

Then the local gauge transformation becomes $U = e^{i\theta + i\boldsymbol{\lambda}\cdot\mathbf{a}}$. In order to make $\psi \to e^{-iq\boldsymbol{\lambda}\cdot\boldsymbol{\phi}(x)/\hbar c}\psi$ locally gauge invariant, the covariant derivative $\mathcal{D}_\mu = \delta_\mu + \frac{iq}{\hbar c}\boldsymbol{\lambda} \cdot \mathbf{A}_\mu$ with eight gauge fields $\mathbf{A}_\mu$ are introduced. Finally, the newly introduced fields that can be identified as gluons need their own propagation terms $\mathcal{L}_g = -\frac{1}{16\pi}\mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu}$, so that the complete QCD Lagrangian shown in Eq. 2.32 results.

$$\mathcal{L} = [i\hbar c\overline{\psi}\gamma^\mu\delta_\mu\psi - mc^2\overline{\psi}\psi] - \frac{1}{16\pi}\mathbf{F}^{\mu\nu} \cdot \mathbf{F}_{\mu\nu} - (q\overline{\psi}\gamma^\mu\boldsymbol{\lambda}\psi) \cdot \mathbf{A}_\mu \tag{2.32}$$

2.1.6 Parton distribution functions

One of the consequences of QCD is that given a collision of two protons, the initial state is ill defined. As any component of a proton can carry any percentage of the actual overall proton momentum at the time of the collision, statistical distributions for the constituents of the protons to interact and to carry which amount of momentum are needed to describe the initial state. These distributions are called parton distribution functions (PDF) $f_i(x, Q^2)$ for partons of type i. The variable x describes the momentum fraction that the specified parton type carries. Lastly, the momentum transfer of the interacting parton is denoted by Q^2.

Because the PDFs are not calculable by perturbative QCD, the parton distribution functions have to be measured at experiments. Many collaborations exist that compile PDF sets and use data from various experiments like HERA, Tevatron or the LHC. For the CMS collaboration, the CTEQ (see Ref. [8]) set of PDFs has been used.

An example of a parton distribution function set of the MSTW collaboration is shown in Fig. 2.2.
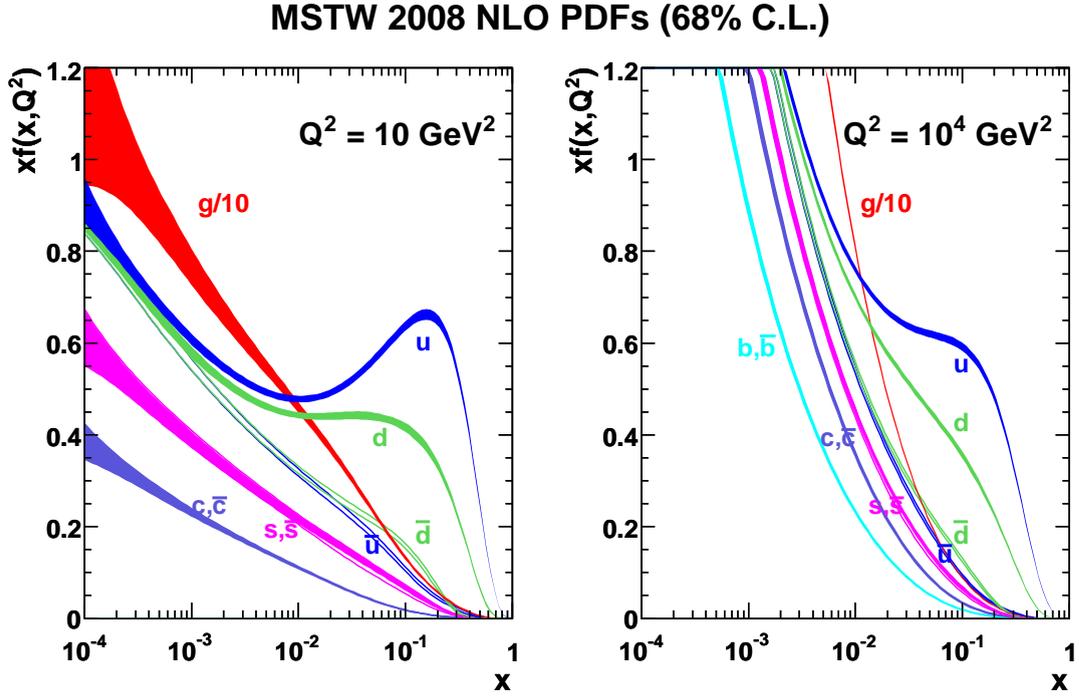

FIGURE 2.2: Taken from Ref. [9], the figure shows that for high momentum fractions, the chance of encountering one of the three valence quarks is highest. At low momentum fractions of the proton, gluons quickly dominate, leading to an abundance of gg processes at the LHC.

2.1.7 Fermi's golden rule

To quantify the likelihood of any interaction between particles to occur, the cross section σ as defined by Fermi's golden rule is used. It states that the cross section is proportional to the squared matrix element $|\mathcal{M}|^2$ and the phasespace divided by a flux factor. The differential cross section for an unambiguous two-body scattering into an area of angles in space, $d\Omega$ is shown in Eq. 2.33

$$\frac{d\sigma}{d\Omega} = (\frac{\hbar c}{8\pi})^2 \frac{|\mathcal{M}|^2}{(E_1 + E_2)^2} \frac{\mathbf{p}_f}{\mathbf{p}_i} \tag{2.33}$$

The matrix element describes the fundamental physics of any interaction and can be calculated from the Lagrangians specified before.

The factor $\frac{1}{(E_1+E_2)^2} \frac{\mathbf{p}_f}{\mathbf{p}_i}$ contains the available energy for the process $(E_1 + E_2)$ and the ratio of either ingoing particle momenta $\mathbf{p}_i$ and either outgoing particle momenta $\mathbf{p}_f$. The more energy is available relative to the needed energy for the process, the larger

the phasespace will be and the more likely the interaction described becomes. Experimentally the cross section is determined using the event rate $\frac{dN}{d\Omega}$ and the luminosity L as specified in Eq. 2.34

$$\frac{d\sigma}{d\Omega} = \frac{dN}{L \cdot d\Omega} \tag{2.34}$$

Therefore, one of the most important quantities at a collider is the integrated luminosity L^{-1} that describes the physical relevance of the recorded data. The expected event rate N before acceptance and selection efficiency is then simply $N = \sigma \cdot L^{-1}$.

2.2 Supersymmetric extensions to the Standard Model

For the introduction and motivation of supersymmetry as a beyond standard model extension, Ref. [7] has been taken as a principle source.

To state a few motivations why extensions of the SM are needed, one can quickly state that neutrinos apparently have mass (see Ref. [10]) and that cosmic microwave background fluctuations and galactic differential rotation speeds hint at the existence of dark matter (see Ref. [11, 12]). Moreover, the expansion speed of the universe measured by studying type Ia supernovae seems to be getting faster which makes it necessary to include a cosmological constant or "dark energy" (see Ref. [13]) or modify the laws of gravity on large scales. Trivially, one can simply state that gravity is not included in the SM.

On the theory side, the SM has the additional problem of large quadratic divergencies for the Higgs field δm_H^2 at a lorentz invariant cutoff scale Λ, as shown in Eq. 2.35.

$$\delta m_H^2 = \frac{12g^2 m_H^2}{2^9 M_W^2 \pi^2} \left(\Lambda^2 - m_H^2 \ln \frac{\Lambda^2}{m_H^2} + \mathcal{O}(\Lambda^{-2}) \right) \tag{2.35}$$

For the SM to hold true at energies up to the Planck scale where gravitational effects would start to matter, radiative corrections via higher order processes need to be contained, in order to let perturbation theory still apply.

So there are several good reasons to search for an extension of the SM. One simple idea is to use the fact that radiative correction terms of fermions and bosons have opposite signs. Therefore, introducing a Majorana spinor operator S that anticommutes with itself and fermionic operators transforms a half-integer spin wavefunction into an integer spin wavefunction and vice versa allows for the definition of an exact copy of each particle with different spin. Doubling the number of particles this way allows to make all divergencies due to radiative corrections moot, because equal and opposite terms of each fermion and their respective bosonic partner cancel. This principle is called supersymmetry and is the only remaining allowed extension of the Poincaré group (containing

translations, rotations and boosts) that constitutes the framework of relativistic quantum field theory.

Of course, this model is inherently too easy to accept, because a supersymmetric partner to a particle (with a prefix "s" for fermions or a suffix "ino" for bosons) should have all the same quantum numbers including mass. A selectron as a spin 1 electron partner should have been found, long ago, if it existed at the same mass scale. Therefore, like for the Z^0 and $W^\pm$ bosons, a spontaneous symmetry breaking mechanism is needed to make superpartners heavy enough to not yet be discovered. In the following section, the minimal supersymmetric standard model (MSSM) will be introduced that contains a lot of simplifications.

2.2.1 MSSM

The MSSM constitutes the most simple version of a SUSY theory. It is defined in the gauge symmetry group $SU(3) \times SU(2) \times U(1)$ and contains a supersymmetric variant of each SM particle with the same quantum numbers, safe spin and mass. The exception is the Higgs field. The MSSM knows not only one Higgs boson, but five $(h^0, \ H^0, \ A^0, \ H^+, H^{\ -})$ by extending the Higgs sector to two doublets $\hat{H}_u$ and $\hat{H}_d$ as shown in Eq. 2.36.

$$\hat{H}_u = \begin{pmatrix} \hat{h}_u^+ \\ \hat{h}_u^0 \end{pmatrix}, \ \hat{H}_d = \begin{pmatrix} \hat{h}_d^- \\ \hat{h}_d^0 \end{pmatrix} \tag{2.36}$$

Unlike the SM, the $\hat{h}_u^0$ vacuum expectation value (VEV) gives mass to up-type quarks, only. The same principle applies to the $\hat{h}_d^0$ VEV giving mass to down-type quarks.

To describe interactions between chiral superfields, a superpotential $\hat{f}$ shown in Eq. 2.37 needs to be defined.

$$\hat{f} = \mu\hat{H}_u^a\hat{H}_{da} + \sum_{i,j=1,3} [(\boldsymbol{f}_u)_{ij}\epsilon_{ab}\hat{Q}_i^a\hat{H}_u^b\hat{U}_j^c + (\boldsymbol{f}_d)_{ij}\hat{Q}_i^a\hat{H}_{da}\hat{D}_j^c + (\boldsymbol{f}_e)_{ij}\hat{L}_i^a\hat{H}_{da}\hat{E}_j^c] \tag{2.37}$$

The full content of superfields and their quantum numbers for QCD $SU(3)_C$, QFD $SU(2)_L$ and QED $U(1)_Y$ is listed in Tab. 2.2.

Field	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
$\hat{L} = \begin{pmatrix} \hat{\nu}_{eL} \\ \hat{e}_L \end{pmatrix}$	1	2	-1
$\hat{E}^c =$	1	1	2
$\hat{Q} = \begin{pmatrix} \hat{u}_L \\ \hat{d}_L \end{pmatrix}$	3	2	1/3
$\hat{U}^c =$	3*	1	-4/3
$\hat{D}^c =$	3*	1	2/3
$\hat{H}_u = \begin{pmatrix} \hat{h}_u^+ \\ \hat{h}_u^0 \end{pmatrix}$	1	2	1
$\hat{H}_d = \begin{pmatrix} \hat{h}_d^- \\ \hat{h}_d^0 \end{pmatrix}$	1	2*	-1

TABLE 2.2: Gauge transformation properties and weak hypercharge assignments for one generation of the MSSM.

While baryon and lepton number are conserved for all renormalizable interactions in the SM, this is not automatically true for the MSSM. One can write the lepton number violating (LNV) terms in Eq. 2.38 and the baryon number violating (BNV) terms in Eq. 2.39.

$$\hat{f}_{LNV} = \sum_{i,j,k} [\lambda_{ijk}\epsilon_{ab}\hat{L}_i^a\hat{L}_j^b\hat{E}_k^c + \lambda'_{ijk}\epsilon_{ab}\hat{L}_i^a\hat{Q}_j^b\hat{D}_k^c] + \sum_i \mu'_i\epsilon_{ab}\hat{L}_i^a\hat{H}_u^b \qquad (2.38)$$

$$\hat{f}_{BNV} = \sum_{i,j,k} \lambda''_{ijk}\hat{U}_i^c\hat{D}_j^c\hat{D}_k^c \qquad (2.39)$$

Since such terms would allow, e.g., the proton to decay, severe experimental limits (see Ref. [14]) exist that suggest that such terms are not realized in nature, unrestrictedly. It has to be said, though, that setting either LNV or BNV couplings to zero is sufficient to eliminate proton decays like the example shown in Fig. 2.3.
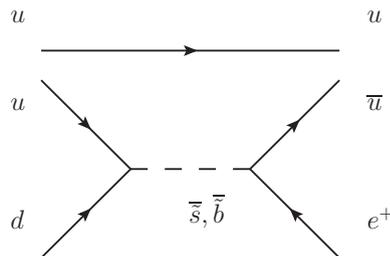

FIGURE 2.3: R-parity violating sparticle-mediated decay of the proton. lepton and baryon number are violated.

Therefore, a new concept of so called "R-parity" can be introduced as defined in Eq. 2.40 to defuse these terms.

$$R = (-1)^{3(B-L)+2s} \tag{2.40}$$

Here, B is the baryon number, L is the lepton number and s is the spin. Given such a concept is realized in nature, a stable lightest supersymmetric particle (LSP) is the consequence that might be a candidate for dark matter, because sparticles conserving R-parity can only be produced in pairs.

The full spontaneous symmetry breaking term $\mathcal{L}_{\text{soft}}$ is shown in Eq. 2.41, 2.42, 2.43, 2.44, 2.45 and 2.46.

$$\mathcal{L}_{\text{soft}} = -[\tilde{Q}_i^\dagger \boldsymbol{m}_{\boldsymbol{Q}ij}^2 \tilde{Q}_j + \tilde{d}_{Ri}^\dagger \boldsymbol{m}_{\boldsymbol{D}ij}^2 \tilde{d}_{Rj} + \tilde{u}_{Ri}^\dagger \boldsymbol{m}_{\boldsymbol{U}ij}^2 \tilde{u}_{Rj} + \tilde{L}_i^\dagger \boldsymbol{m}_{\boldsymbol{L}ij}^2 \tilde{L}_j + \tilde{e}_{Ri}^\dagger \boldsymbol{m}_{\boldsymbol{E}ij}^2 \tilde{e}_{Rj} + m_{H_u}^2 |H_u|^2 + m_{H_d}^2 |H_d|^2] \tag{2.41}$$

$$- \frac{1}{2}[M_1 \overline{\lambda}_0 \lambda_0 + M_2 \overline{\lambda}_A \lambda_A + M_3 \overline{\tilde{g}}_b \tilde{g}_b] \tag{2.42}$$

$$- \frac{i}{2}[M_1' \overline{\lambda}_0 \gamma_5 \lambda_0 + M_2' \overline{\lambda}_A \gamma_5 \lambda_A + M_3' \overline{\tilde{g}}_b \gamma_5 \tilde{g}_b] \tag{2.43}$$

$$+ [(\boldsymbol{a_u})_{ij} \epsilon_{ab} \tilde{Q}_i^a H_u^b \tilde{u}_{Rj}^\dagger + (\boldsymbol{a_d})_{ij} \tilde{Q}_i^a H_{da} \tilde{d}_{Rj}^\dagger + (\boldsymbol{a_e})_{ij} \tilde{L}_i^a H_{da} \tilde{e}_{Rj}^\dagger + \text{h.c.}] \tag{2.44}$$

$$+ [(\boldsymbol{c_u})_{ij} \epsilon_{ab} \tilde{Q}_i^a H_u^b \tilde{u}_{Rj}^\dagger + (\boldsymbol{c_d})_{ij} \tilde{Q}_i^a H_{da} \tilde{d}_{Rj}^\dagger + (\boldsymbol{c_e})_{ij} \tilde{L}_i^a H_{da} \tilde{e}_{Rj}^\dagger + \text{h.c.}] \tag{2.45}$$

$$+ [b H_u^a H_{da} + \text{h.c.}] \tag{2.46}$$

The 3×3 complex matrices $\boldsymbol{a}$ and $\boldsymbol{c}$ describe trilinear scalar interactions. Furthermore, the mass of the gauginos is described by six parameters (M_i, M_i'). When counting all parameters in the MSSM, one can say that it contains nine parameters out of its gauge sector, five parameters due to its Higgs sector, and 207 parameters in the flavor sector. If one removes parameters that can be dropped by field redefinitions, still 178 parameters remain.

Since such a number of parameters is not manageable in terms of exploration at the present time, typically several approximations are done to reduce the number of parameters allowed. The most effective simplification is to heed third generation Yukawa couplings, only. This effectively reduces the 207 flavor sector parameters to 23. Usually, it is also required that the SUSY breaking parameters and μ are of the order of the weak scale, so at most a few TeV, because otherwise the main advantage of SUSY to fix the radiative corrections to the Higgs mechanism breaks down. This very approximation becomes prominent in the next section as a motivation for searching for third generation sparticles.

2.2.2 Motivation for searching for third generation sparticles

Because the attractivity of SUSY as a theory relies heavily on radiative corrections to the Higgs mass canceling, the current limits on colored particles mostly exceed the range (see Fig. 2.4) in which this canceling would happen without severe theoretical modifications (or "finetuning") discouraging SUSY searches. A notable exception to this trend are third generation sparticles that have a lower production cross secton, as shown in Fig. 2.5.
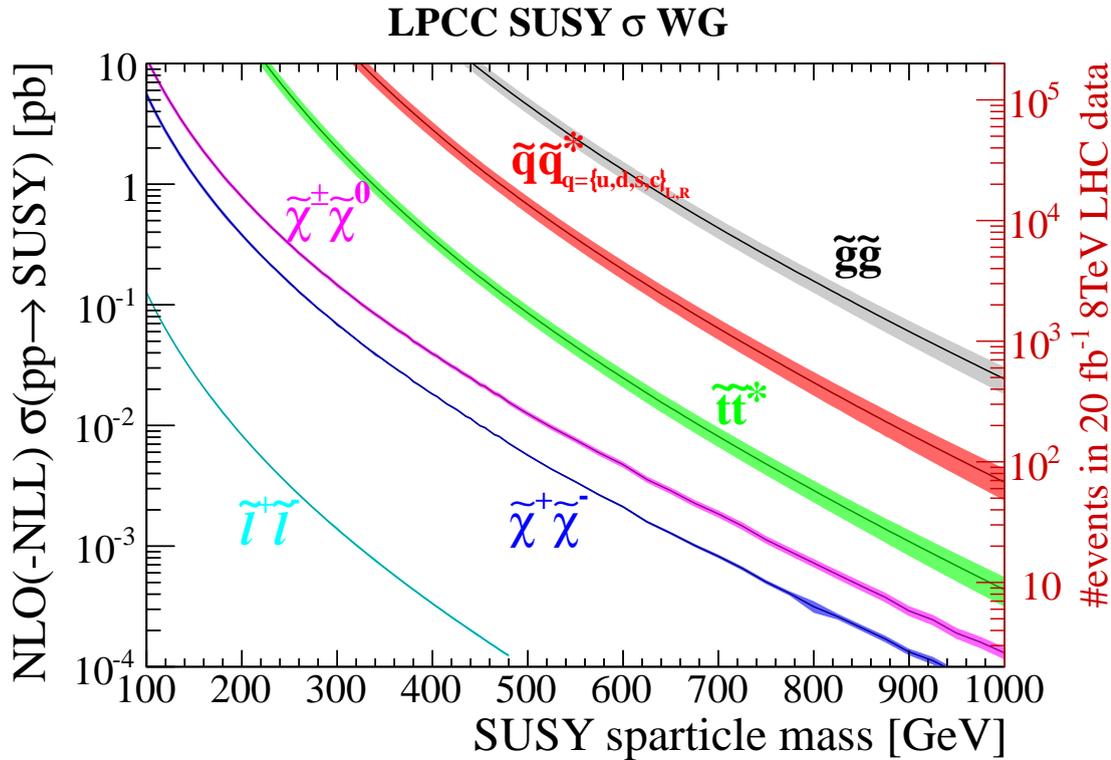
FIGURE 2.5: Taken from Ref. [16], documented in Ref. [17], the cross sections for different sparticle production processes for $\sqrt{s} = 8$ TeV are shown. Since the production cross sections for electroweak and third generation sparticles are generally lower than for gluinos or light squarks, limits could only be set accordingly.

As the largest contribution to the radiative Higgs correction stem from the top quark, having the stop quark at a reasonable mass would already cancel the most problematic radiative terms.

Moreover, results from cosmology regarding the abundance of dark matter in the universe put constraints on the type of sparticles allowed and the lightest stable supersymmetric particle that could, given R-parity holds, not decay to any other particle. It would
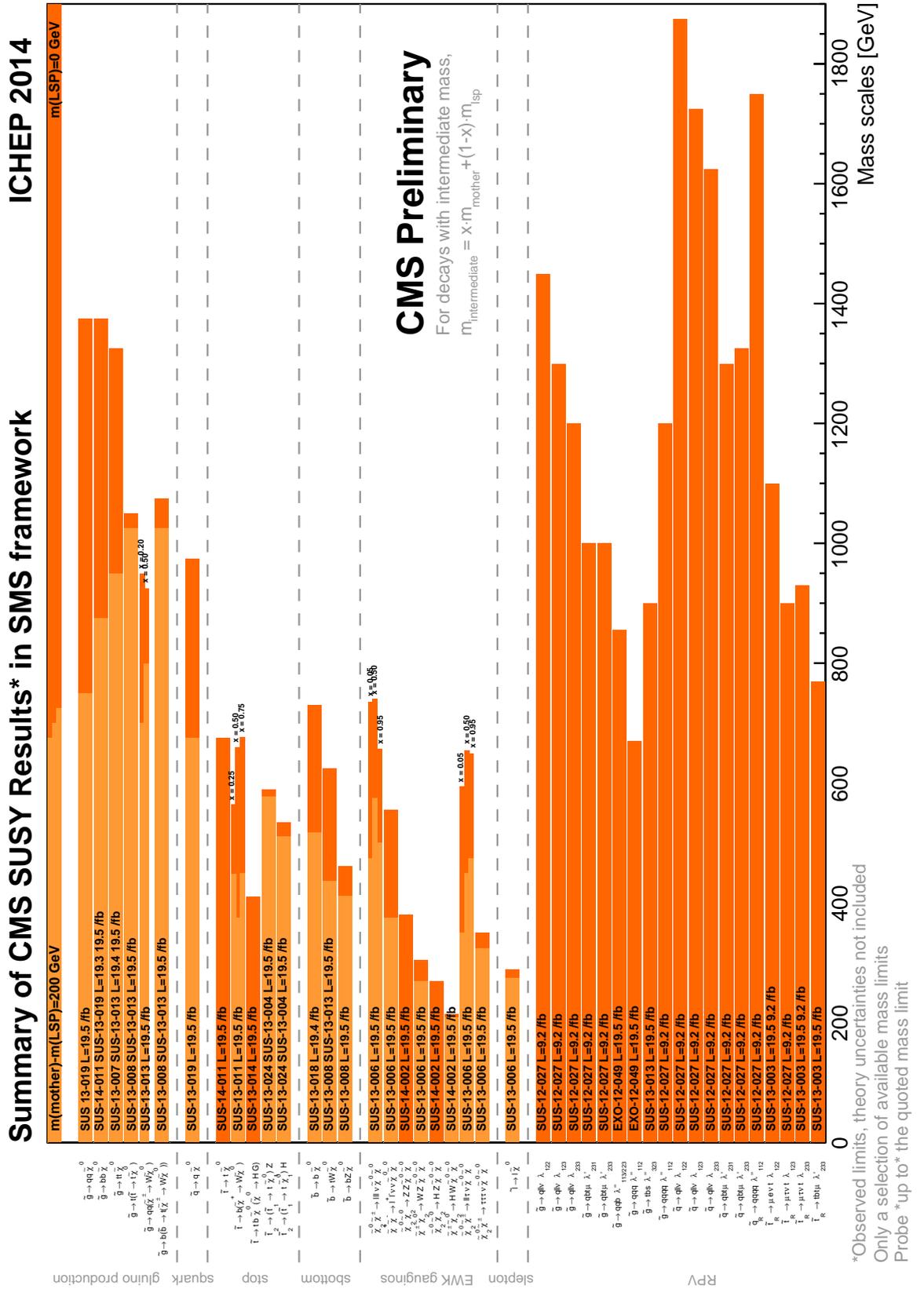
FIGURE 2.4: Taken from Ref. [15], the plot shows the current limits on different sparticles in diverse production channels. For colored sparticles, the limits usually exceed 1 TeV, with the notable exception of third generation sparticles.

therefore be an ideal candidate for dark matter, but in turn it would need to fulfill some
very specific requirements, as discussed in the following section.

2.2.3   Motivation for searching for compressed spectra SUSY

The relic dark matter density in the universe is measured by cosmological experiments.
If SUSY is responsible for producing dark matter, then the SUSY parameters give rise
to the relic density in cosmology. In Fig. 2.6, it is shown which region of phasespace
would be favored by such assumptions for staus as third generation sparticles.
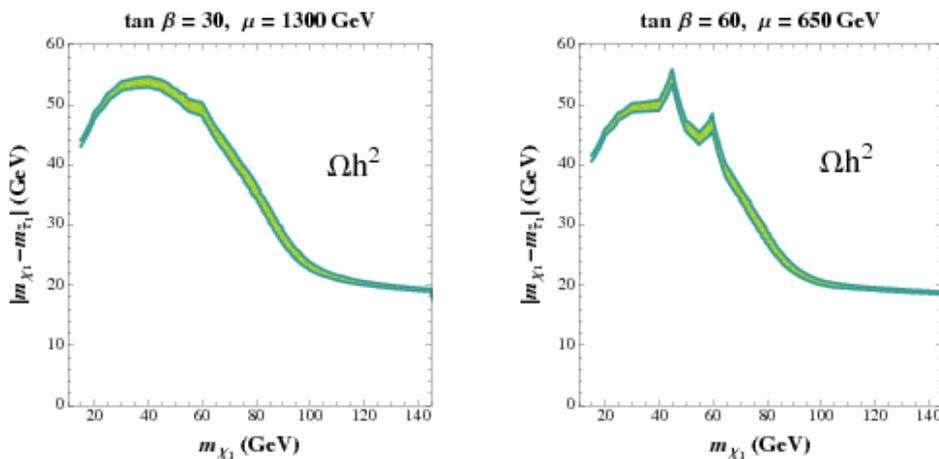
FIGURE 2.6: Taken from Ref. [18], the figure shows the allowed mass-difference
between the lightest chargino and the stau for various chargino masses. In
this case, "allowed" means that the resulting relic dark matter density in the
universe would fit with observations with cosmology. Please note that the
mass-differences are, in general, small.

It can be concluded that staus should be looked into, because this is one of the few
hints for physics beyond the SM that is very specific. The problem is that in the
decay $\tilde{\chi}_1^\pm \rightarrow \tilde{\tau}\nu_\tau \rightarrow \tilde{\chi}_1^0\tau\nu_\tau$, the mass-difference translates into τ momentum. As these
differences are predicted to be small, the τ have to be soft in most cases. When all
accessible sparticles are close together in terms of mass, the spectrum is denoted as
"compressed".

In the experiment, τ are one of the most challenging objects to use and soft $\tau^\pm \rightarrow \pi^\pm\nu_\tau$
decays are indistinguishable from soft proton proton scattering processes. Therefore,
they cannot be triggered and this region is experimentally hard to access, as shown in
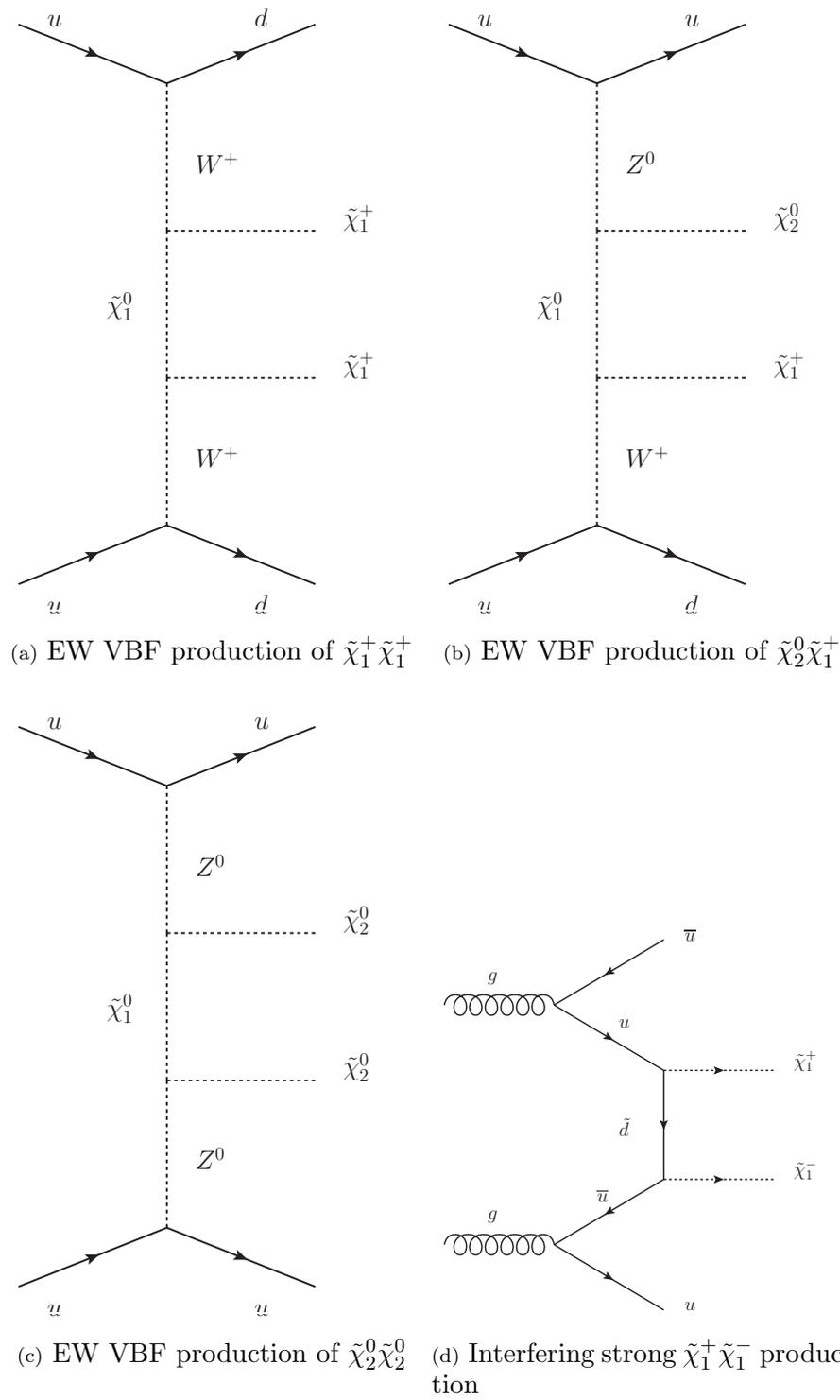Fig. 2.7.

FIGURE 2.7: Taken from Ref. [19], the limits on triple stau production are shown. Theoretically, the region up to the diagonal where the $\tilde{\chi}_1^0$ would cease to be the LSP, is allowed.

As can be seen, most of the region close to the diagonal that would be disallowed by the $\tilde{\chi}_1^0$ ceasing to be the lightest particle is yet unexplored and would be favored by cosmological restraints as seen in Fig. 2.6. This conclusion makes alternative approaches to cover this region worthwhile to pursue, as demonstrated in the following section.

2.2.4 Vector boson fusion channel SUSY searches

As the electroweak production cross section is very low (see Fig. 2.5), alternative production modes are sought that allow for a differentiation from backgrounds and for triggering. Vector boson fusion production modes feature two jets with a high rapidity gap and at least two leptons in the final state, if the chargino and neutralino to sleptons and/or leptons branching fractions are high. Example feynman diagrams of VBF electroweak production processes with two charginos, a neutralino and a chargino or two neutralinos as well as an example of the strong production with an intermediate squark that interferes with the electroweak production is shown in Fig. 2.8.

(a) EW VBF production of $\tilde{\chi}_1^+\tilde{\chi}_1^+$ (b) EW VBF production of $\tilde{\chi}_2^0\tilde{\chi}_1^+$

(c) EW VBF production of $\tilde{\chi}_2^0\tilde{\chi}_2^0$ (d) Interfering strong $\tilde{\chi}_1^+\tilde{\chi}_1^-$ production

FIGURE 2.8: Fig. a/b/c) show examples for VBF production. Fig. d) shows an example of strong chargino production that interferes with the electroweak production.

It is yet disputed whether the interference is constructive or destructive, but an effect on all VBF production ross sections shown is expected.

While the leptons are typically soft, as expected from compressed spectra, the jets have high momenta, go into opposite detector hemispheres at typically high pseudorapidities and do not balance each other. This allows for discrimination by characterization of the forward jets, as shown in Fig. 2.9. In the recently published Ref. [1], this principle has
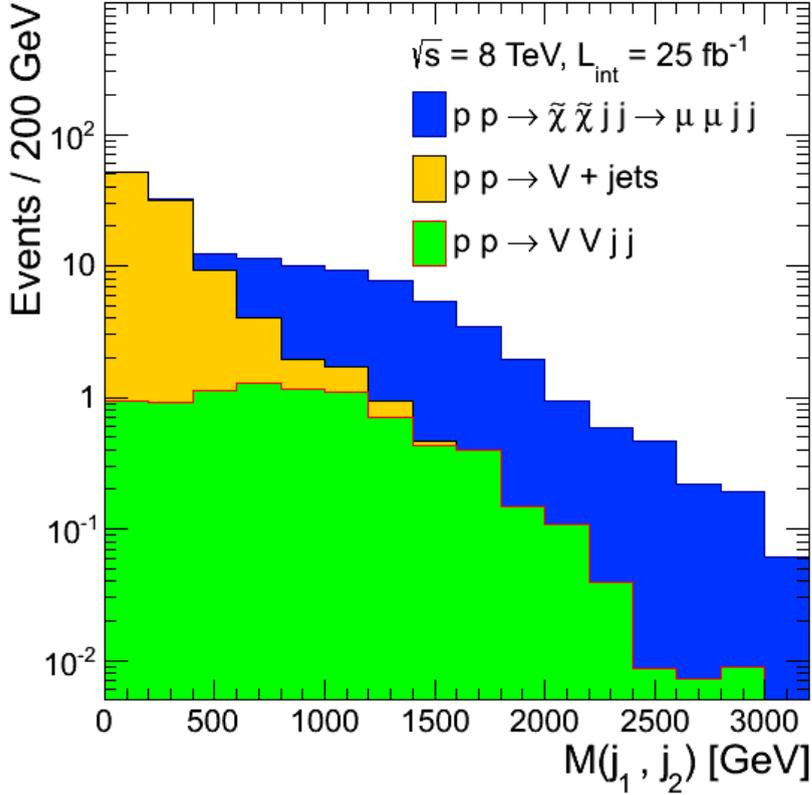
FIGURE 2.9: Taken from Ref. [20], background and signal processes are shown in a stacked plot of the di-jet invariant mass. It can be seen that there is discriminative power in using such a variable.

been used to search for compressed spectrum SUSY in search channels characterized by the pair of produced central leptons that stem from the $\tilde{\tau} \to \tau \tilde{\chi}_1^0 \to \nu_\tau + e\nu_e \vee \mu\nu_\mu \vee \pi$ decay chain. In total, the $\mu^\pm\mu^\pm$, $\mu^\pm\mu^\mp$, $\mu^\pm\tau_h^\pm$, $\mu^\pm\tau_h^\mp$, $\mu^\pm e^\pm$, $\mu^\pm e^\mp$, $\tau_h^\pm\tau_h^\pm$, and $\tau_h^\pm\tau_h^\mp$ channels contributed to the exclusion limits shown in Fig. 2.10. The general strategy for the VBF searches has been to search for third generation sleptons well-motivated by radiative Higgs corrections and dark matter relic density. The searches have been performed in exclusive decays of electroweak production modes to $\tilde{\tau}$ in as many decay channels of the resulting τ as viable. For most analyses, one μ in the event allowed triggering with very soft leptons, while for the hadronic channels with the highest branching ratio only a τ_h trigger on hard τ_h was available. In this thesis, the contribution of the author to the $\tau_h^\pm\tau_h^\pm$ channel will be discussed in detail in Chap. 6.
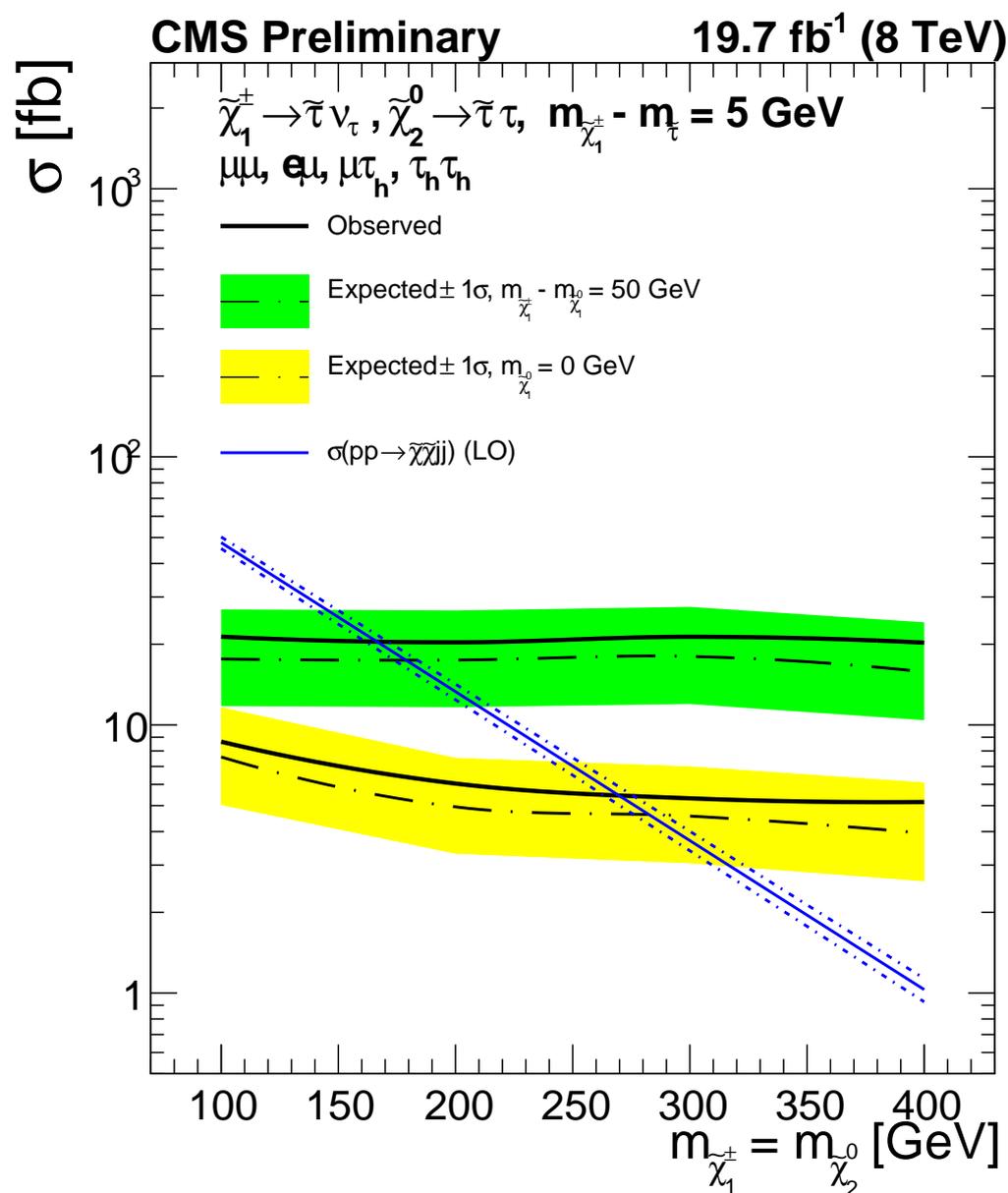
FIGURE 2.10: Taken from Ref. [1], the combined limit on compressed spectrum (green) and negligible mass LSP (yellow) SUSY by the VBF SUSY searches is shown. The compressed region up to $M(\tilde{\chi}_0^\pm) < 170$ GeV is excluded

Chapter 3

Experimental setup

In this chapter, the Large Hadron Collider supplying accelerated proton bunches for collisions and the Compact Muon Solenoid experiment used to detect and reconstruct the resulting collision events are described.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) (see Ref. [21]) is the largest machine constructed by man to date and also the most powerful particle accelerator currently available. It is situated at the European Organisation for Nuclear Research (CERN) compound in Geneva, Switzerland. The accelerator can be used for proton or heavy ion beam acceleration.

All data used in this thesis was recorded in proton-proton collisions durnig the 2012 data taking. The specifications of the beams in 2012 and the design values of the LHC are specified in Tab. 3.1.

Parameter	Value in 2012	Design value
Beam energy [TeV]	4	7
β^* in IP 1,2,5,8 [m]	0.6, 3.0, 0.6, 3.0	0.55
Bunch spacing [ns]	50	25
Number of bunches	1374	2808
Average bunch intensity [protons per bunch]	$1.6\text{-}1.7 \cdot 10^{11}$	$1.15 \cdot 10^{11}$
Normalized emittance at start of fill [mm.mrad]	2.5	3.75
Peak luminosity [cm^{-2}s^{-1}]	$7.7 \cdot 10^{33}$	10^{34}
Max. mean number of events per bunch crossing	≈ 40	19
Stored beam energy [MJ]	≈ 140	362

TABLE 3.1: LHC specifications reached in 2012 compared to the design values, taken from Ref. [22].

The beam acceleration is performed by radiofrequency cavities operated at 400.8 MHz. The circumference of the LHC is 27 km. 1232 superconducting niobium-titanium dipole magnets force the proton beams to stay on course. The magnets are cooled by superfluid helium and operate at 4.8 T. Quadrupole, sextupole, and octupole magnets are in use to collimate the beams.

The two LHC beams are collided at four interaction points where the four main experiments are located. One of the main experiments is ALICE (see Ref. [23]), a detector with a time projection chamber optimized for heavy ion collision detection. Another main experiment is ATLAS (see Ref. [24, 25]), a multipurpose detector experiment. Also, there is the CMS experiment (see Ref. [26, 27]), the second multipurpose detector at the LHC that will be described in detail in Sec. 3.2. Finally, there is LHCb (see Ref. [28]), an asymmetric detector optimized for measurements of b-quark hadrons.

A sketch of the LHC, its preaccelerators of the respective experiments is provided in Fig. 3.1.

FIGURE 3.1: Taken from Ref. [29], the figure shows the general layout of the LHC, it's preaccelerators and connected experiments.

3.1.1 2012 run at $\sqrt{s} = 8$ TeV

The high luminosity delivered by the LHC gives rise to challenges and opportunities. In terms of opportunities, the integrated luminosity acquired during the 2012 run allows for probing new physics and measuring standard model properties to high precision (e.g. the top quark mass, see Ref. [30]). The data recorded over time is shown in Fig. 3.2.
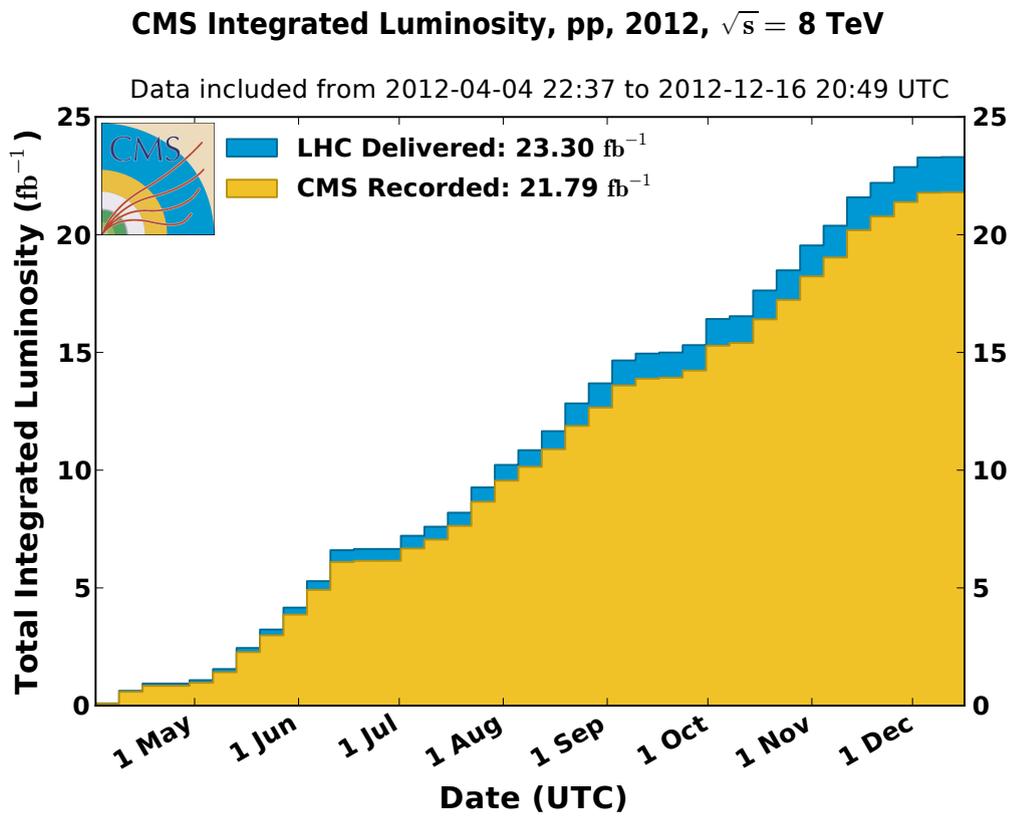
FIGURE 3.2: Taken from Ref. [31], the figure shows the integrated luminosity delivered by the LHC and recorded by the Compact Muon Solenoid experiment during 2012.

The high rate of interactions also challenges the LHC experiments substantially. Especially, triggering events of interest is problematic as most collisions lead to minimum bias events. Minimum bias events are expected at a rate of $\mathcal{O}(20)$ MHz and need to be reduced to a rate that is manageable by data acquisition and reconstruction at roughly $\mathcal{O}(100)$ kHz (see Ref. [32]) before a further reduction to $\mathcal{O}(400)$ Hz for final storage tages place.

Moreover, even events of interest are polluted by pileup. In 2012, 21 additional vertices on average could be reconstructed per event as shown in Fig. 3.3.
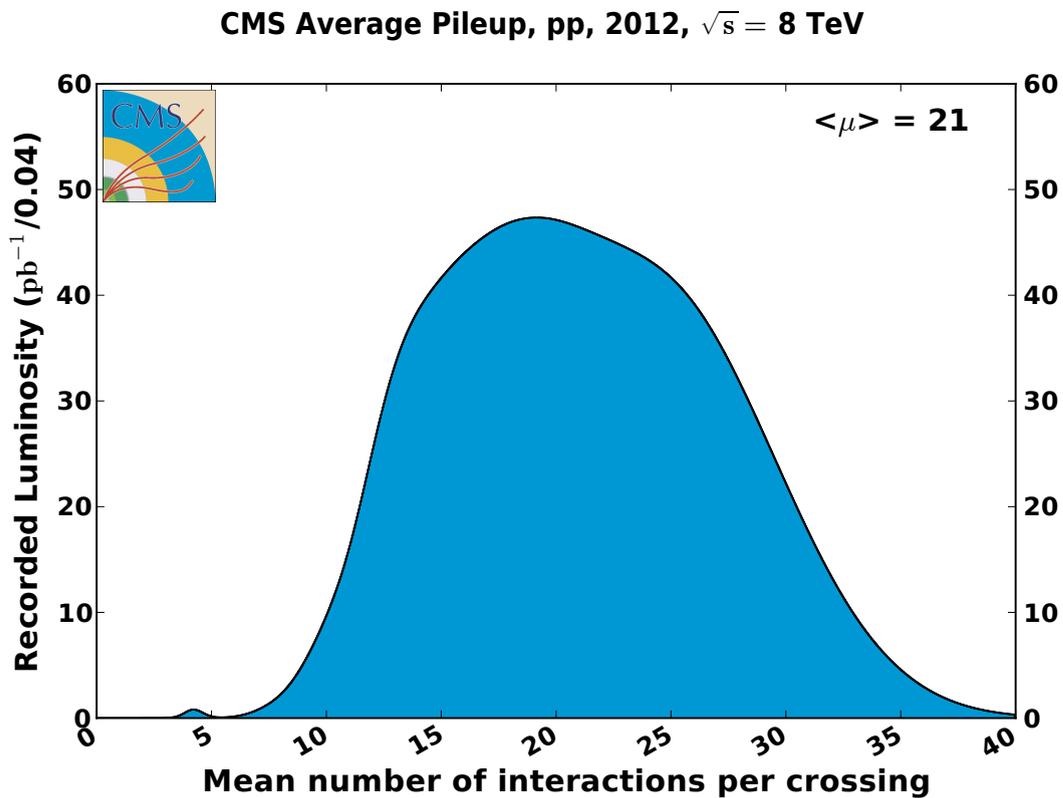
FIGURE 3.3: Taken from Ref. [31], the figure shows the mean number of interactions per crossing at the Compact Muon Solenoid experiment during 2012 run conditions.

3.2 Compact Muon Solenoid detector

All information in this section and it's subsections is taken from Ref. [33], if not otherwise specified.

The Compact Muon Solenoid (CMS) detector is a general purpose detector at the LHC. It is built such that it fits within the tight spatial restrictions given by the size of the access tunnel to the experimental cavity, as it was built aboveground and later lowered into the cavity. The complete detector is shown in Fig. 3.4 and has a size of 28.7 m length times 15 m diameter in a cylindrical shape.
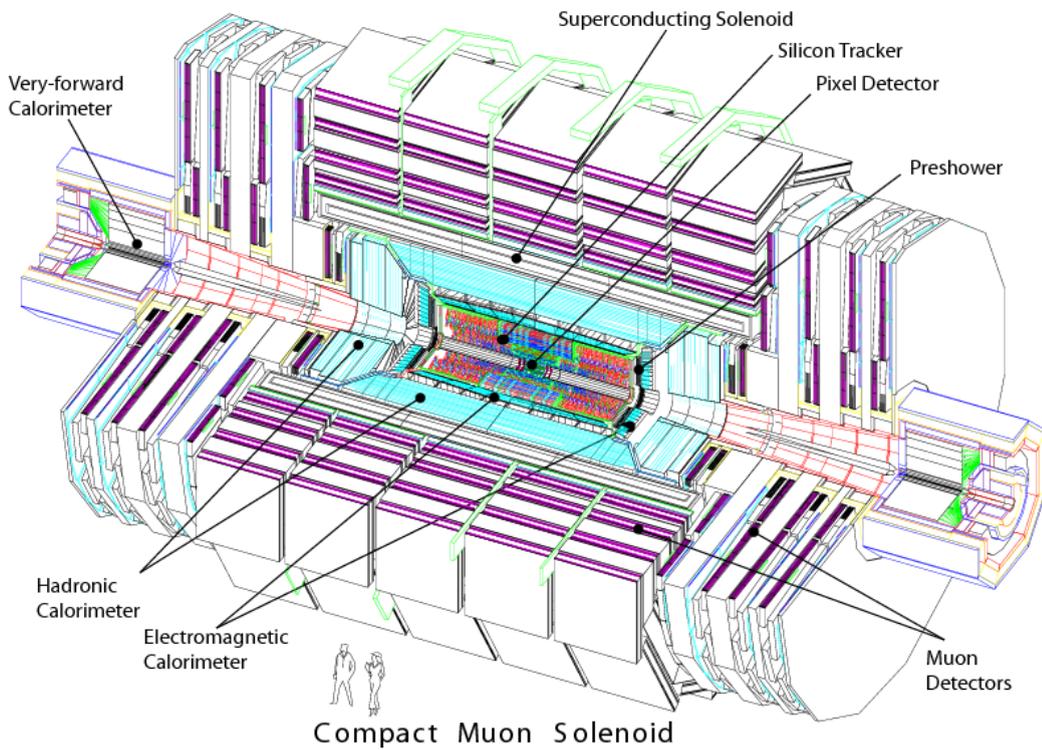
FIGURE 3.4: Taken from Ref. [33], the figure shows a cross section of the CMS detector. From the interaction point in the center of the detector, subdetectors are built in cylinder symmetry around the beamline. Closest to the interaction point is the silicon pixel tracker, followed by the silicon strip tracker, then the electromagnetic calorimeter, the hadronic calorimeter, the superconducting coil and finally the muon systems.

Due to the compact size of the detector, it has to be very dense, in order to absorb most particles produced at the interaction point. Therefore, it weighs approximately 14 Mg. The detector coordinate system is set up in cylinder coordinates with the center of origin at the nominal interaction point. The x axis points to the center of the LHC ring, the y axis upwards and the z axis anticklockwise with respect to the beam direction. The polar angle θ is defined in the plane spanned by the y-z plane from the positive z axis. The azimuthal angle ϕ is defined in the x-y plane from the positive x axis. Moreover, the polar angle is usually expressed in terms of the pseudorapidity $\eta = -\ln(\tan(\theta/2))$ that is equal to the lorentz-invariant angle called "rapidity" for massless particles (see Ref. [34] on use cases for this definition).

The most commonly used observables are transverse energy and transverse momentum of physics objects. The definition of "transverse" in this case is the projection of any object's four-vector on the x-y plane. As the beam at the interaction point has no

momentum transverse to the beamline, but the initial longitudinal momentum of inter-
acting partons is unknown, only transverse momentum components are balanced such
that their vectorial sum is zero, if all components were reconstructed perfectly.

3.2.1   Tracker

Closest to the beamline is the silicon pixel detector. It consists of three barrel layers at
4.4 cm, 7.3 cm and 10.2 cm radial distance to the center of the beamline, respectively,
and two pixel disks. The pixel tracker resolution is roughly 10 μm in r-ϕ and about 20 μm
in z. This resolution is mainly needed for primary and secondary vertex identification.
In total, the pixel tracker has 66 million pixels.

Located at $200 < r < 1100$ mm, the silicon strip detector contains ten barrel and twelve
disc layers with the innermost two layers of both systems being dual layers affixiated to
each others' back at an angle of 100 mrad. The barrel has 9.6 million strips and consists
of the Tracker Inner Barrel (TIB), Tracker Outer Barrel (TOB), Tracker Endcap (TEC)
and Tracker Inner Discs (TID) subsystems.

The complete tracker design is shown in Fig. 3.5.
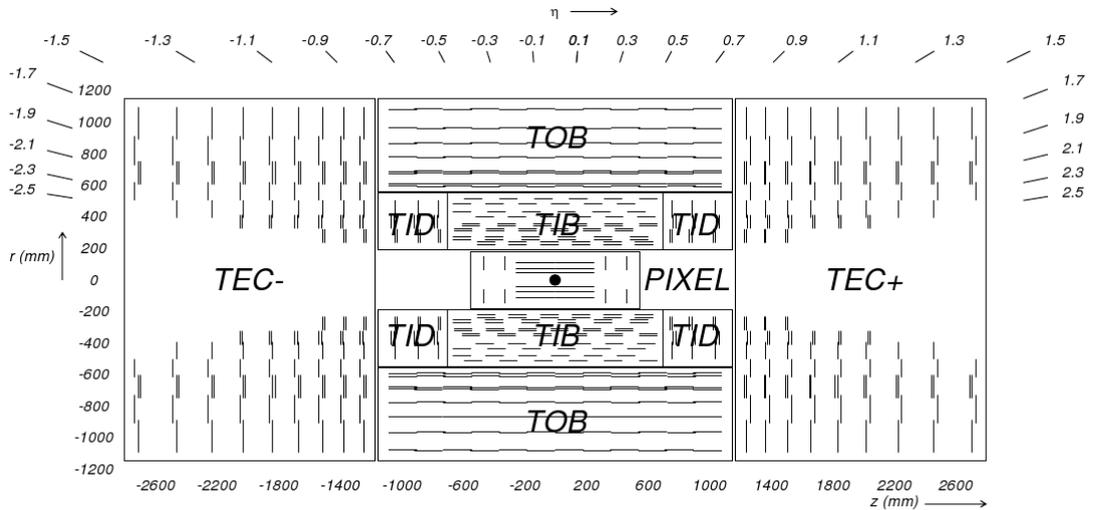
FIGURE 3.5: Taken from Ref. [35], the figure shows the CMS tracking system.
The pixel detector and the different parts of the strip detector are shown with
their relative positions in the detector. Please note that the tracker coverage
ends at $|\eta| > 2.5$.

The tracking system is held by carbon-fibre structural supports and contained in a
support tube under temperature control.

3.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) consists out of 75848 lead tungstate ($PbWO_4$) crystals (61200 for the barrel part (EB) and 14648 for the endcap parts (EE)). These homogeneous crystals are absorbers and scintillators and are instrumented by avalanche photodiodes (APD) that detect scintillation light. The radiation length X_0 at which traversing electrons lose e^{-1} of their initial energy on average due to bremsstrahlung is 89 mm for such a crystal. The crystals in the CMS detector have a length of 22 cm in the endcaps and 23 cm in the barrel section. That length also marks the expected onset of electromagnetic showers in the ECAL.

The crystals are built into a support structure in a quasi-projective way by being tilted by 3° with respect to the interaction point in η and ϕ with an overall granularity of $0.0174\,[\eta] \cdot 0.0174\,[\phi]$.

The detection rate of photons in the APDs depends on the temperature, making it necessary to use a cooling system. Furthermore, the rate of photons detected depends on the transparency of the crystals. As the crystals become opaque by radiation damage, the transparency is constantly measured by a system of lasers, in order to determine calibration constants as shown in Fig. 3.6.
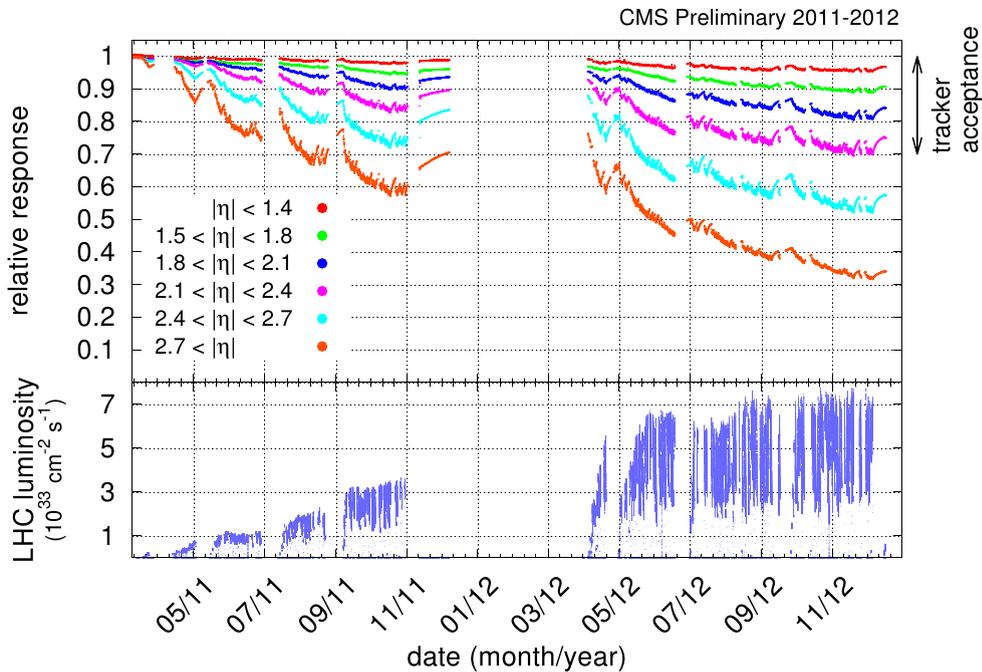
FIGURE 3.6: Taken from Ref. [36], the figure shows the laser calibration constants determined during the $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV runs. During irradiation, the ECAL crystals become more opaque. Given sufficient time without irradiation, a recovery of transparency can be observed.

Finally, the energy (E) resolution (σ) of the ECAL for a photon or electron is shown in Eq. 3.1.

$$\frac{\sigma}{E} = \sqrt{\left(\frac{2.83\%\sqrt{\text{MeV}}}{\sqrt{E}}\right)^2 + \left(\frac{124\ \text{MeV}}{E}\right)^2 + (0.26\%)^2} \tag{3.1}$$

3.2.3 Hadronic calorimeter

The hadronic calorimeter (HCAL) is a sampling calorimeter. There are three main sections of the HCAL. It is divided into the barrel (HB) in $|\eta| \leq 1.4$, the endcaps (HE) in $1.4 < |\eta| \leq 3.0$, and the forward calorimeters (HF) in $3.0 < |\eta| \leq 5.174$.

The HB contains brass absorbers and plastic scintillators in turns with the exception of the first and last absorption layers being steel. The scintillating material is connected to APDs by wavelength-shifting fibres. The HE is made in a similar arrangement.

The HF consists out of iron and quartz-fiber calorimeters. The fibers are embedded into a steel absorber and generate Cherenkov light when exposed to charged particles. The light is then detected by photomultipliers. A distinction between electromagnetic and hadronic showers is possible by using short fibers ranging from the middle to the end of the steel absorber together with long fibers that go through the whole absorber structure. Only the long fibers are supposed to detect the electromagnetic parts of the shower.

In terms of overall structure, the HCAL consists of towers in $0.175\ [\eta] \cdot 0.175\ [\phi]$ for the HF and $0.087\ [\eta] \cdot 0.087\ [\phi]$ for the HB and HE in an arrangement of squares fitting on top of 5^2 ECAL crystals. The towers can be interpreted as a coverage in $\Delta\eta$ and the corresponding spacing is used for the choice of binning edges for the jet energy corrections, as described in Tab. 5.5.

A general problem of the HCAL is its low energy resolution described in Eq. 3.2 according to Ref. [37].

$$\frac{\sigma}{E} = \frac{115.3\%\sqrt{\text{GeV}}}{\sqrt{E}} + 5.5\% \tag{3.2}$$

The response of electromagnetic shower components is far better due to the high quality homogeneous ECAL at CMS than the response of hadronic showers components. Therefore, any calorimeter tower is undercompensating (the ratio of electron response e to hadron response h is $\frac{e}{h} > 1$), leading to non-gaussian resolution tails and an overall resolution depending on jet energy and jet composition according to Ref. [38].

3.2.4 Superconducting solenoid

For correct charge determination and a good resolution of momentum measurements of charged particles by the tracker, a powerful magnetic field of 3.8 T is employed. It is

generated by a superconducting solenoid with a length of 12.9 m and a diameter of 5.9 m. The solenoid is made from 2168 windings of Niobium-titanium (NiTi) cables and stores a magnetic field energy of 2 GJ generated by a current of 18 kA. The main design feature besides the strong magnetic field is that tracker and calorimetry are contained inside of the solenoid.

3.2.5 Muon system

In almost all cases, charged particles passing through the calorimeters and solenoid unabsorbed are muons. In order to unambiguously identify these muons and measure their momentum, a series of drift trubes and resistive plate chambers ($|\eta| < 1.2$) and a similar arrangement of cathode strip chambers ($0.9 \leq |\eta| < 2.4$) and resistive plate chambers ($|\eta| < 1.6$) are employed. These systems are interspersed by iron return yokes that are installed to contain the magnetic field.

3.2.6 Trigger system

According to Tab. 3.1, the design value of bunch spacings for the LHC is 25 ns. This corresponds to a bunch crossing rate of 40 MHz.

Hardware processing restrictions allow for a data transfer rate to storage elements at roughly $\mathcal{O}(1 \text{ kHz})$. That necessitates the use of a trigger system to reduce the number of events to be stored by five orders of magnitude.

At the CMS experiment, a two-level trigger is employed. Therefore, the level one (L1) trigger at the hardware level has to be very efficient. It is built to reduce the rate by identification of signatures in the detector hardware by three orders of magnitude.

Afterwards, the raw detector data is processed by a server farm with simplified reconstruction algorithms and subject to further sets of selections, the so called "High-level trigger" (HLT) as the level two of the trigger system. Events passing both stages of the trigger system are written to data storage for immediate or for later reconstruction (parked).

Chapter 4

Physics objects

Any physical particle that is detectable needs to be reconstructed, directly or indirectly, by taking information from one (or several) part(s) of the detector and applying algorithms to identify it. In this chapter, event simulation that bridges the gap between the theoretical formulation of particles to the reconstructed objects detected is introduced. Furthermore, the different objects like jets, τ or missing transverse energy that are specifically used in this thesis will be explained.

4.1 Event generation and simulation

In order to compare predicted physics processes modeled by theory sources to data, the collisions and the physics processes happening thereafter need to be simulated.

The first step is to determine the particles in a collision that are the main facilitators of the following process. Parton distribution functions (PDFs) provide the necessary information. For the "Summer12" production of the CMS collaboration that was used in this analysis for simulated events, the CTEQ 6.6L PDFs Ref. [8] are used. Differences to other sets of PDFs are considered for deriving systematic uncertainties.

The second step after the incoming particles are determined is to calculate the matrix element and therefore the probability of the physics processes happening. Available generators with this functionality are e.g. MadGraph (see Sec. 4.1.1), POWHEG (see Sec. 4.1.2), Pythia6 (see Sec. 4.1.3) or Herwig++ (see Sec. 4.1.4).

There are three different kinds of parton interactions that have to be distinguished. The hard process describes the most energetic interaction in an event that has triggered the detector trigger system. In contrast, the underlying event (UE) encompasses all low-energetic parton interactions that happened due to the interaction of different parts of the same protons that gave rise to the hard process. Finally, pileup (PU) is the

sum of all radiation due to low-energetic parton interactions of protons other than the ones giving rise to the hard process. PU can happen before, during and after the event of interest, as long as the resulting detector hits still lie within the time window of the detector electronics and are therefore not directly distinguishable from the event of interest, itself.

After a hard process has been calculated, the outgoing particles need to decay and, in the case of QCD related processes, fragment and hadronize. The latter is called a parton shower and can be performed by e.g. Herwig++ or Pythia6.

Particles reaching the detector material are then subject to a detector simulation by Geant4 (see Sec. 4.1.5) or a fast but simplified version by Fastsim (see Sec. 4.1.6).

4.1.1   MadGraph

MadGraph v4 Ref. [39] is a matrix element (ME) generator. It contains next to leading order (NLO) perturbative calculations and is needed to be used, in order to get more accurate descriptions of more complicated processes than $2 \to 2$. Simulated samples using MadGraph are denoted as such by a simple "madgraph" in their descriptive name. Such samples include e.g. Z+jet simulated background samples.

Since MadGraph does not include a parton shower (PS) simulation, a matching between the ME part provided by MadGraph and the PS part provided by a PS generator like Herwig or Pythia is needed. In the CMS collaboration, the so called "MLM" matchingscheme is used where partons are clustered with a k_T algorithm (see Sec. 4.2.1) and the angular distance ΔR between partons and their respective showers is calculated. If the transverse momentum of any parton shower is below a cutoff or the angular distance in a k_T metric is above a certain cutoff, the whole event is rejected. An exception is done for high multiplicity samples where extra jets are allowed.

The number of additional partons that can be simulated by MadGraph depends on the simulated process. For the samples used in this thesis, a maximum of four additional partons was simulated.

4.1.2   POWHEG

POWHEG Ref. [40] is a matrix element generator that contains NLO perturbative calculations. Simulated samples using POWHEG are labeled by "powheg" in their description. Among those samples are e.g. Z→ll simulated samples.

4.1.3 Pythia6

Pythia6 Ref. [41] is a composite program that contains a simple $2 \rightarrow 2$ matrix element generator, as well as a hadronization and fragmentation simulation. For QCD simulated samples, the matrix element generation is used in the samples used in this thesis. In such a case, the sample will be labeled by a "pythia6" without the addition of another matrix element generator like MadGraph or POWHEG in front.

A generator tune is needed in order to describe UE contributions to an event, because out of first principles, this is not calculable. Instead, events are recorded without a hard process and only the residual activity in the detector due to underlying event is measured.

Here, the tune Z2* Ref. [42] has been used in combination with Pythia6. It has been found to describe the behaviour of the data best out of a set of tunes tested in Ref. [42].

4.1.4 Herwig++

Herwig++ Ref. [43] contains a $2 \rightarrow 2$ matrix element generator and is able to perform fragmentation and hadronization. Together with the tune LHC-UE-EE-3-CTEQ6L1 Ref. [44] (denoted by "TuneEE3C" in sample names), it is used as a crosscheck to Pythia6.

4.1.5 Geant4

For all officially produced non-signal datasets and therefore for all simulated background samples used in this thesis, the Geant4 Ref. [45] framework is used in order to simulate the interaction of particles with the CMS detector. It contains all interactions with the detector material, nuclear interactions and shower simulation in the calorimeters or even before them.

4.1.6 Fastsim

Fast simulation Ref. [46] is used to generate simulated, e.g., SUSY signal events in vast quantities at reasonable time scales. It simplifies the shower simulation in the calorimeters towards lateral and transverse shower function templates and neglects details of detector-particle interactions, in order to reduce processing times.

4.2   Jets

The confinement of QCD processes disallows free quarks from being observed. Instead, a multitude of coloured particles ending in color-neutral bound states is produced in a process called hadronization.

Given hadronization happens to a massive particle at rest with respect to the frame of reference, an observer would see a spherical shower of particles (mostly of pions). But the particles produced in the hard interactions of pp events are not at rest. They propagate at large fractions of the speed of light, leading to a different shape of the observed phenomenon. In the direction of propagation of the hadronizing particle, a cone-like spray of particles evolves, until it is absorbed in the detector. This phenomenon needed some time to be observed and to be theoretically described, because at the first occurences at high energy physics colliders (e.g., at DORIS) particle multiplicities were low and opening angles very wide due to low energies of the hadronizing particles according to Ref. [47].

Today, we are at the verge of the next era of jets, where energies are sufficient for several jets being observed as a single jet object Ref. [48].

4.2.1   Jet reconstruction

Having an access to the particles hadronizing into jets is a complicated procedure. Observables like energy and direction of the originating particle need to be well described. In order to have a theoretically and experimentally equally viable way of unique determination of jets, sequential reconstruction algorithms are the current tool in use. They are infrared safe, meaning that additional soft radiation from, e.g., underlying event or pileup do not bias jet energy or jet direction. Moreover, they are collinear safe. This means that a jet should change neither direction nor energy reconstructed, if one more or one less collinear particle splitting takes place.

These algorithms Ref. [49] use a distance argument (see Eq. 4.1) and a distance cutoff (see Eq. 4.2) in a variable metric, in order to recombine the four vectors of energy clusters (see Sec. 4.2.2) to jet objects. The necessary algorithms are implemented in the Fastjet library (see Ref. [50]) and realized by generating Voronoi diagrams of all clusters considered, always recombining the closest two in the chosen metric, if the distance is below the cutoff distance, until no cluster remains for consideration.

$$d_{ij} = \min(p_{T,\,i}^{2n}, p_{T,\,j}^{2n}) \frac{dR_{ij}^2}{d_{\max}^2} \tag{4.1}$$

$$d_{iB} = p_{T,\,\mathrm{i}}^{2n} \tag{4.2}$$

There are different established metrics that can be used for jet clustering and that have different phenomenologies and cases of use. The Cambridge-Aachen algorithm with $n = 0$ uses a purely geometric argument without considering cluster energies and is useful for jet substructure algorithms of very collimated high energetic particle decays. The k_t algorithm with $n = 1$ recombines soft particles first and assures in lepton colliders without soft energy backgrounds that all energy is clustered.

In this thesis, the standard jets are clustered by the anti-k_t Ref. [51] algorithm that uses $n = -1$ and a default $d_{max} = 0.5$, called "ak5" in shorthand. It supports very circular jet shapes and centers around large energy deposits. Therefore, it is ideally suited for reconstructing jets with high particle multiplicity and small jet opening angles in the high pileup environment at the LHC. Due to the strictly circular structure, some energy is typically lost (called "out of cone" showering), but only the offset energy due to diffuse energy by underlying event and/or pileup of a well-defined jet area is added as a surplus.

4.2.2 Particle Flow

The CMS detector's layout features very good tracking and it supports a powerful ECAL. On the downside, the calorimetry is heavily undercompensating according to Ref. [26]. That means that the response of pions is lower than the response of electrons, due to the powerful homogeneous ECAL. As the electromagnetic and hadronic energy fractions of jets vary and the response depends on the aforementioned energy fractions, such dependencies worsen the jet energy resolution (JER).

For a general calorimeter-based measurement, the JER as defined in Eq. 4.3 applies according to Ref. [52].

$$\frac{\sigma(E)}{\langle E \rangle} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \qquad (4.3)$$

Here, a is a stochastic term, b a term for electronic noise and c a constant term due to detector effects and miscalibrations. Typical values for CMS would be $a = 1.25\sqrt{\text{GeV}}$, $b = 5.6$ GeV and $c = 3.3\%$, according to Ref. [53].

The solution of the CMS collaboration to this problem for the measurement of jet energies is to use the Particle Flow (PF) algorithm Ref. [54]. It basically combines information from all subdetectors to identify different classes of particles. Charged particles that comprise $\approx 65\%$ of the energy of a jet on average are measured in the tracker and in the calorimeter, separately, then matched. A weighting procedure ensures that wherever the measurement of a particle momentum with the tracker is more accurate (up to 100 GeV particle transverse momentum), it is taken into account with a larger weight than the calorimeter measurement, improving the overall precision.

In general, the PF algorithm differentiates between muons, electrons, photons, charged

hadrons and neutral hadrons. In the aforementioned order, linked energy deposits in the calorimeters and their corresponding tracks are removed from further consideration. Only if no track is left pointing to a calorimeter energy deposition after substracting all the linked energies, the rest is considered to originate from neutral particle(s) (photons in the ECAL and neutral hadrons in the HCAL).

During this process, the calorimeter energy resolutions are taken into account. Still, the finite granularity of the calorimeters and no longitudinal segmentation at all give rise to new uncertainties on the PF jet energy resolution. Energy deposits and their link to tracks are subject to possible confusion of which track is responsible for which energy deposit, especially if the same tower in the ECAL and the HCAL is lit. Furthermore, showers can leak from one tower into the next, making it hard to correctly link the track pointing to the tower of origin to the additional energy deposits in adjacent towers, because the shower shape itself is not observed.

To summarize, the particle flow algorithm is used at CMS for improving the energy resolution of measurements of charged components of jets by using the tracker to at least partially take over the measurement.

4.2.3 Charged hadron substraction

In order to suppress the influence of pileup on events, tracks and their associated calorimeter energies from all but the most energetic vertex can be removed. This process is called charged hadron substraction (CHS). After the application of this algorithm on particle flow, the resulting PFCHS procedure is the currently recommended source of clusters for jet clustering in the CMS collaboration.

4.2.4 Jet identification

Some quality criteria are recommended to be required (see Ref. [55]), in order not to misidentify objects primarily originating from colorless particles like photons or electrons as jets or enrich detector noise. In Tab. 4.1, those requirements are summarized for the loose working point used for all purposes in this thesis.

PF jet ID	loose
neutral hadron fraction	< 0.99
neutral EM fraction	< 0.99
number of constituents	> 1
muon fraction	< 0.8
charged EM fraction	< 0.99
$|\eta| < 2.4$	
charged hadron fraction	> 0
charged multiplicity	> 0
charged EM fraction	< 0.99

TABLE 4.1: Quality requirements on jets for the loose working point. Inside tracker coverage ($|\eta| < 2.4$), additional requirements are applied. "EM" is shorthand for "electromagnetic".

4.2.5 b-tagging

In order to identify jets originating from bottom quarks, the non-negligible lifetime of bottom quarks before hadronization occurs can be used, because for light quarks hadronization starts basically instantaneously. The reconstruction of a secondary vertex relative to the primary vertex allows for a relatively good identification rate and purity. Several algorithms for b-tagging are available at CMS Ref. [56]. In this thesis, the combined secondary vertex b-jet tags are used with a discriminator requirement > 0.244 (loose working point). B-tagging is only applied in order to veto events containing top quarks that have a $> 99\%$ likelihood of decaying further to a bottom quark and leptons or jets.

4.3 Electrons

According to Ref. [57], the electron identification at CMS for non-isolated electrons works by a tracker driven seeding. This means that a silicon strip track matched to a pixel seed (at least two pixel hits compatible with the primary vertex) is used as a starting point for the reconstruction of an electron. The trajectories are reconstructed by usage of a gaussian sum filter.

Furthermore, electrons are absorbed in the ECAL and the energy losses therein due to bremsstrahlung have a distinct shape. Such a shape is called a "supercluster" and geometrically matched to the tracker seed.

Then, a preselection is applied that is based on multivariate methods with inputs being

the agreement between the track and ECAL energy measurements, supercluster shape and track quality criteria. In this thesis, the "version 5" of this multivariate variant is used at the medium working point in order to discriminate against hadronically decaying τ.

4.4 Muons

In this thesis, a loose muon identification criterium of "version 3" is used to discriminate against hadronically decaying τ. According to Ref. [58], the loose identification means that the muon in question is required to be a PF muon and that it is also considered to be a tracker and/or global muon.
In Ref. [59], these different categories are defined as:

global μ A track in the muon system is successfully combined with a track reconstructed in the silicon tracker by the Kalman-filter technique.

tracker μ A track with $p_T > 0.5$ GeV and $p > 2.5$ GeV is extrapolated to the muon system taking into account possible course deviations and the magnetic field. At the end of the extrapolation, at least one muon chamber hit has to be found.

PF μ Depending on the relative amount of energy deposited in the vicinity of the muon candidate, flexible selection criteria optimized on reconstructing muons inside jets are applied.

To discriminate against charged hadrons penetrating the HCAL (but still leading to energy depositions in the HCAL), isolation requirements are set by summing up the energy of all PF neutral and charged candidates in a cone of $\Delta R = 0.4$ around the muon candidate.
Furthermore, the muon identification requires at least one pixel hit, at least ten tracker hits, at least one hit in the muon system, the global μ criteria, some Kalman-filter output quality criteria and a relative isolation.

4.5 Taus

In this thesis, hadronically decaying one-pronged $\tau^\pm \to \pi^\pm \nu_\tau$ (τ_{h}) are used. Three-pronged decays ($\tau^\pm \to \pi^\pm \pi^\mp \pi^\pm \nu_\tau$) are omitted due to a requirement of one-pronged decays at HLT level for the trigger used in the VBF analysis presented in Chap. 6. Therefore, three-pronged decays are not used for any purpose in this thesis.

In order to identify τ_h, a hadron plus strip (HPS) combined isolation algorithm is employed Ref. [60]. PF charged hadrons and photons are taken into consideration and requirements on the particle multiplicity and invariant mass of charged hadrons and neutral pions in a cone of size $\Delta R = \frac{2.8}{p_T}$ are made. Moreover, a larger cone with size $\Delta R = 0.5$ around the τ_h candidate is taken into account for isolation. Three different working points (loose, medium, and tight) are available for this purpose and all three are used in this thesis. The labels and requirements on τ_h objects in this thesis are summarized in Tab. 4.2.

label	requirement
againstElectronMediumMVA5	> 0.5
againstMuonLoose3	> 0.5
leadPFChargedHadrCand_pt	≥ 5.0 GeV
decayModeFindingNewDMs	> 0.5
signalPFChargedHadrCands_size	$= 1$
& one of the following:	
byTightIsolationMVA3newDMwLT	≤ 0.5
byMediumIsolationMVA3newDMwLT	≤ 0.5
byLooseIsolationMVA3newDMwLT	≤ 0.5

TABLE 4.2: The first requirement is a veto against electrons that requires a failure of a medium working point of the electron ID (see Sec. 4.3). The second requirement is a similar veto against muons requiring a failure of a loose working point (see Sec. 4.4). Then, a charged hadron PF candidate with at least 5 GeV transverse momentum is required, as well as a confirmed one-pronged decay. Finally, one out of three different isolation discriminant working points is required.

Expected τ_h efficiencies are $\approx 50\%$ as average across all possible decay channels at the loose working point and $\approx 25\%$ at the tight working point, according to Ref. [61]. In the same reference, the problem of the τ_h energy scale is raised. Part of the τ_h is not reconstructed, as it is carried away by the ν_τ out of the decay. For all purposes in this thesis, no attempts at using a τ_h energy scale correction or employing methods using the missing transverse energy are used to correct for this effect.

4.6 Missing transverse energy

The missing transverse energy ($\vec{\not{E}}_\mathrm{T}$) is defined as the imbalance of transverse momentum of all visible particles in an event Ref. [62]. If all visible components were perfectly

measured, the resulting $\vec{\displaystyle{\not}E}_{\mathrm{T}}$ would then be the transverse momentum carried by invisible particles like neutrinos or, e.g., $\tilde{\chi}_1^0$ for SUSY scenarios.

In the following, only PF $\vec{\displaystyle{\not}E}_{\mathrm{T}}$ will be explained, as it has been used frequently in this thesis. The $\vec{\displaystyle{\not}E}_{\mathrm{T}}$ is then defined as the negative vectorial sum of transverse momenta of all PF particles. In most cases, $\displaystyle{\not}E_{\mathrm{T}}$ as the magnitude of $\vec{\displaystyle{\not}E}_{\mathrm{T}}$ is used in this context.

4.6.1 Type I corrected MET

As the type I corrected missing transverse energy for all work in this thesis, a definition has been used which propagates the jet energy corrections (see Chapter 5) towards the $\vec{\displaystyle{\not}E}_{\mathrm{T}}$. This procedure is undertaken in order to minimize a bias on the $\vec{\displaystyle{\not}E}_{\mathrm{T}}$ from jet miscalibrations.

The propagation on simulated events is shown in Eq. 4.4 and on data in Eq. 4.5.

$$\vec{\displaystyle{\not}E}_{\mathrm{T}} = \vec{\displaystyle{\not}E}_{\mathrm{T}}^{\mathrm{raw}} + \sum_i \vec{p}_T(\mathrm{jet}_i) \cdot (L1 \cdot L2L3 - L1) \tag{4.4}$$

$$\vec{\displaystyle{\not}E}_{\mathrm{T}} = \vec{\displaystyle{\not}E}_{\mathrm{T}}^{\mathrm{raw}} + \sum_i \vec{p}_T(\mathrm{jet}_i) \cdot (L1res \cdot L2L3 \cdot L2L3res - L1res) \tag{4.5}$$

When a different definition of a type I correction is used, it is mentioned, separately.

4.7 Filters against unphysical events

The CMS collaboration recommends a number of filters to be used against events caused not by hard proton-proton interactions, but by miscellaneous effects. This section lists the recommended filters and their general motivations.

- HBHE noise filter: The Hadronic Barrel Hadronic Endcaps (HBHE) noise filter is used to remove noise originating from hybrid photodiode (HPD) or readoutbox (RBX) related noise. Pulse shape variables are used to distinguish real energy depositions from noise. Additionally, large energy deposits in the HCAL around HPDs or RBXs not matched to a multitude of tracks and/or a sufficiently large counterpart of ECAL energy deposits are removed.

- CSC Tight Halo Filter: Residual gas inside the LHC vacuum chamber or limiting apertures sometimes collide with the LHC beams. Such collisions lead to secondary particle production. Charged particles may be deflected by the beam optics and build a "halo" that can interact with the detector. In order to protect against such events, a cathode strip chamber (CSC) based level one trigger is used that takes into account a typical path in η followed by beam halo muons.

- HCAL laser event filter: The HCAL can be calibrated with a laser beam. If the laser beam firing coincides with a bunch crossing, the laser energy deposit can be misinterpreted as physical energy deposit due to the event. In order to remove these events, this filter removes a list of all events during which the laser calibration was active.

- ECAL dead cell filter: Some ECAL cells are known to be inactive. If such an ECAL cell is hit by a reconstructed jet, part of the jet cannot be reconstructed, leading to missing transverse energy. In order to protect against such events, a geometrical distance argument between any jet in an event and any dead cell is used for the filter.

- Tracking failure filter: In rare cases, sizeable calorimeter deposits without matching tracks are observed. Either too many tracker seeds lead to a failure to reconstruct individual tracks or the hard interaction happened too far away from the nominal interaction point for proper reconstruction. To discriminate against such events, the sum of momenta of all tracks in the event is compared to the sum of momenta of all jets in the event. The filter requires at least one good vertex to be in the event and at least 10 % of the energy in the event to be measured by the tracker.

- EE bad supercrystal filter: Two 5×5 ECAL crystal arrangements (supercrystal) in the detector regularly yield anomalously high energies. This filter removes events with too high energies in either of the aforementioned supercrystals.

- ecal Laser Corr Filter: In rare cases, the laser calibration of ECAL crystals can fail. If this happens, unphysically large correction factors (> 3 in the barrel and > 8 in the endcaps as of 2012) may be in the database. This filter checks for large correction factors and removes events in which these are applied.

- trk POG filters: Anomalous events with partly or fully aborted track reconstruction or strip tracker coherent noise were identified. This filter removes events without any tracks and events where the number of pixel tracker and strip tracker hits deviate largely.

Chapter 5

Jet energy correction

Jets are the most common high-level objects to be used at a hadron collider. The classical challenge for jet reconstruction is to reliably and accurately reconstruct jet direction and jet energy. As explained in Sec. 4.2, jets at CMS are composite objects clustered out of particle flow (PF, see Sec. 4.2.2) objects that, by themselves, contain direction and energy information. These jets are called PF jets.

Each PF object does have some intrinsic calibration and thus, each detector component, be it ECAL (see Sec. 3.2.2), HCAL (see Sec. 3.2.3) or tracker (see Sec. 3.2.1) is also precalibrated. Still, the combination of such precalibrated objects does not amount to a full calibration for the composite object. This has many reasons. While the ECAL is continuously calibrated, the HCAL is not. Another reason is that part of the jet energy is absorbed by the detector material and can therefore not be reconstructed. Lastly, all detector components have noise suppression thresholds that depend on the radiation damages and the pileup environment. So even if energy is registered at a certain point of the detector, that does not completely assure its consideration for jet reconstruction, if aforementioned energy falls below a threshold. Thus, there needs to be a dedicated jet energy calibration.

In order to perform such a calibration, there needs to be an observable for the quality of the calibration. Therefore, a response distribution needs to be defined, where the response $\mathcal{R}$ is an observable relating the reconstructed energy of the jet to the energy of the particles that formed the jet. As the LHC is a proton proton collider, the initial longitudinal energies of the particles involved in the hard scattering process are unknown, only observables transverse to the beam direction are known to have a vectorial sum of zero, given the event can be fully reconstructed. For jets, it is customary to use the transverse momentum p_T of the jets as an observable. A general formulation of a

response using the jet transverse momentum is shown in Eq. 5.1.

$$\mathcal{R} = \frac{p_T^{\text{probe}}}{p_T^{\text{reference}}} \tag{5.1}$$

Examples for reference objects can be generator particles or jets in simulation or well-reconstructable objects like $Z \to \mu\mu$ balancing a jet.

An example of a distribution of a response is shown in Fig. 5.1.
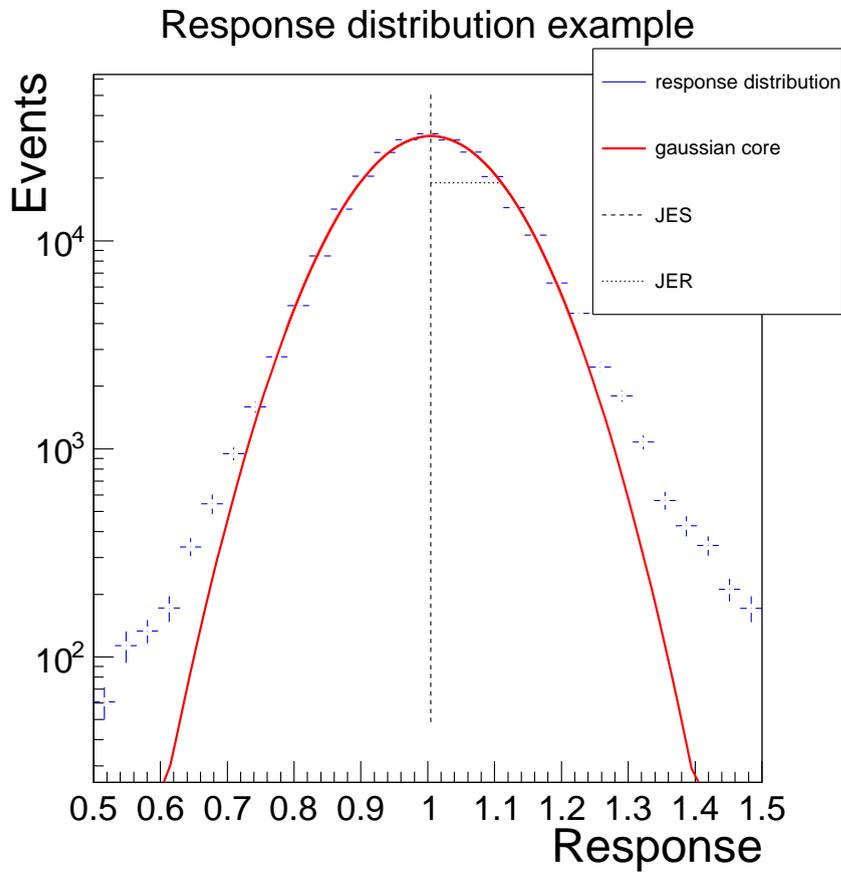
FIGURE 5.1: A realistic example of a response distribution. The mean of the gaussian core of the response is the jet energy scale (JES) and the width of this gaussian core is the jet energy resolution (JER).

The ultimate goal is to, on the one hand, have a sensible comparison of data and simulation by assuring that the jet energy scale (JES) is at the exact same level for both. On the other hand, the shape of the jet energy resolution (JER) should also be as similar as possible in data and in simulation.

5.1 CMS factorized jet energy correction scheme

In order to calibrate the JES to the same level for data and simulation, the CMS experiment uses a scheme of several factorized corrections that are applied in a fixed order to account for different physical effects in simulation and data. Each of these corrections has a shortscript name containing **L** for level, a number indicating the position in the factorized scheme and a suffix **res** (or the lack thereof) indicating a residual correction (or a simulation based correction). All non-residual corrections are determined in simualation and applied on data and simulation, in order to correct the JES in simulation to generator jet level. The residual corrections are determined after the non-residual corrections and compare data to simulation, in order to correct the JES in data to the corrected simulation level, effectively making generator jets and reconstructed jets in data comparable. The full scheme, including all corrections and variables taken into account is shown in Fig. 5.2.
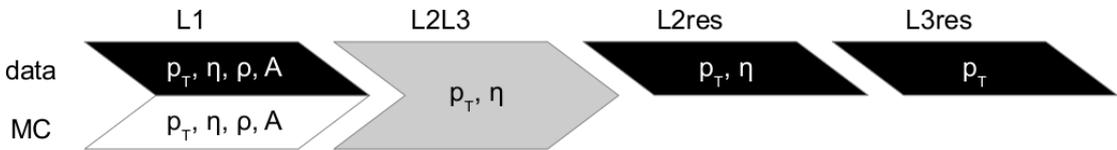
FIGURE 5.2: Illustration of the CMS jet energy scale correction scheme. Black parts are applied on data, white parts on simulation and grey parts on data and simulation, equally.

First, the L1 correction removes additional contributions to the reconstructed energy by pileup and underlying event contributions, as they are indistinguishable. Then, the L2L3 correction corrects the reconstructed jet energy without pileup back to the generator level in simulation as a function of jet transverse momentum and jet direction in terms of pseudorapidity of the jet axis. Those are all the necessary steps for simulated events.

For data, the L1res correction determines scaling factors with respect to the L1 pileup correction out of data, in order to account for differences between the pileup modeling in simulation and the pileup observed in data. Then, the L2res and L3res correct for the differences of the JES in data and simulation that remain after L1resL2L3 corrections in terms of pseudorapidity and transverse momentum of the jets. After the application of L1L2L3 corrections on simulated events and L1resL2L3L2resL3res corrections on data, the JES is similar for both datasets, meaning that the JES in data has been shifted such that it is comparable to the generator jet energy scale in simulation.

Finally, the JER in simulation is broadened by scale factors, in order to make the gaussian core width similar to the one measured in data.

5.2 L1: Pileup offset correction

Taken from Ref. [63], additional proton-proton collisions coinciding during the same bunch crossing as events of interest add diffuse radiation that is partly registered in the tracker and that deposits energy in the calorimeters. Such contributions are called "in-time pileup" (IT PU). A second source of pileup is "out-of-time pileup" (OOT PU). It contributes energy measured in the calorimeters from previous or subsequent collisions that cause an overlap of their energy deposition registration with the calorimeter energy integration times during the event of interest.

While charged IT PU contributions can be reduced using charged hadron substraction (CHS, see Sec. 4.2.3) and OOT PU is reduced by careful choice of the signal integration time of the calorimeter cells, some energy contributions remain. This diffuse additional energy is also clustered into jets and therefore biases the JES.

The pileup offset correction takes into account the number of good quality primary vertices N_{PV} and the diffuse offset energy density ρ per event.

Two simulated samples are used, in order to estimate the offset energy for each jet as a function of A, ρ, η_{jet} and $p_{T\,\mathrm{jet}}$. Both samples contain the same hard interactions, with one sample having simulated pileup and the other sample not having any pileup simulation. The jets of both samples are matched within $\Delta R < 0.5 \cdot R_{\mathrm{cone}}$ to each other and their common particle jet with similar transverse momentum. This way, the average offset energy is determined and the L1 correction is generated that substracts this additional energy from each jet. An illustration of the offset as a function of different μ denoting the poissonian mean of the pileup vertices is shown in Fig. 5.3
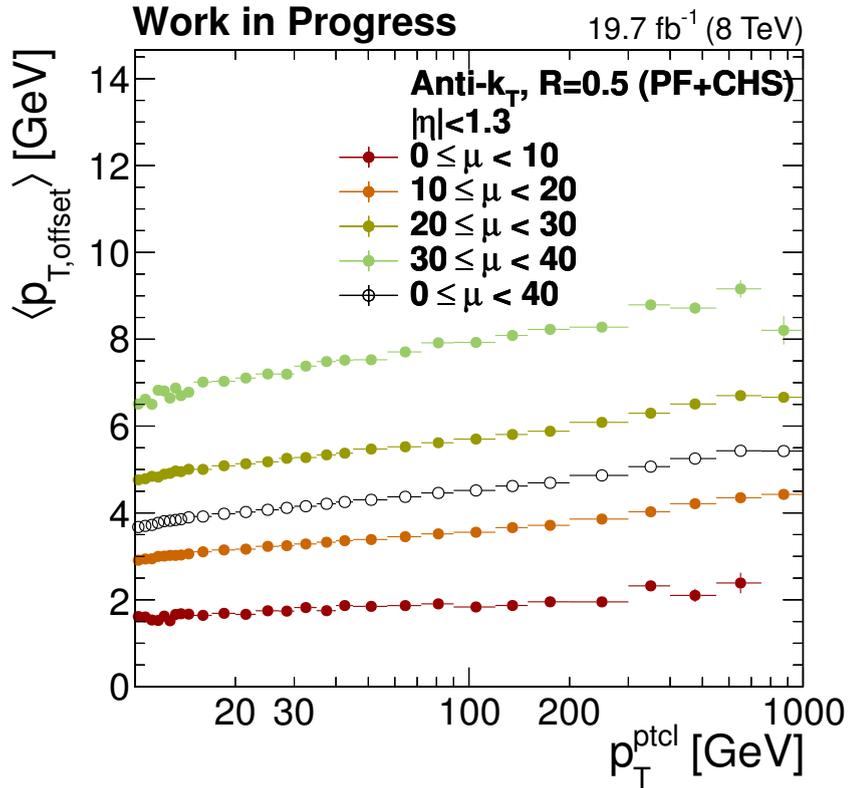
FIGURE 5.3: From Ref. [63], the offset energy by pileup per generator transverse momentum for ak5 PF+CHS jets for different poissonian means μ of the number of vertices per event in the barrel region. The substraction of the offset energy shown is, essentially, the L1 correction.

5.3   L2L3: Simulated response corrections

Taken from Ref. [63], the detector simulation (see Sec. 4.1 and 4.1.5) describes various detector effects. Additionally, fragmentation of gluons and quarks, parton showers and underlying event activity are described by Pythia6 tune Z2* (see Sec. 4.1.3). Both, detector effects and particle fragmentation have an effect on the JES.

The L2L3 correction derives the JES as the mean of the MC truth response, as defined in Eq. 5.2, thereby correcting for particle fragmentation and detector effects to the level described by simulation.

$$R(p_{T,\,\text{ptcl}}, \eta) = \frac{\langle p_{T,\,\text{reco}} \rangle}{\langle p_{T,\,\text{ptcl}} \rangle}[p_{T,\,\text{ptcl}}, \eta] \qquad (5.2)$$

The binning is denoted as $[p_{T,\,\text{ptcl}}, \eta]$, so a binning in particle level transverse momentum and reconstructed pseudorapidity. A matching between reconstructed and particle level is performed within $\Delta R < 0.5 \cdot R_{\text{cone}}$.

The jet response that is to be corrected by the L2L3 correction is shown as a function of η and $p_{T,\,ptcl}$ in Fig. 5.4.
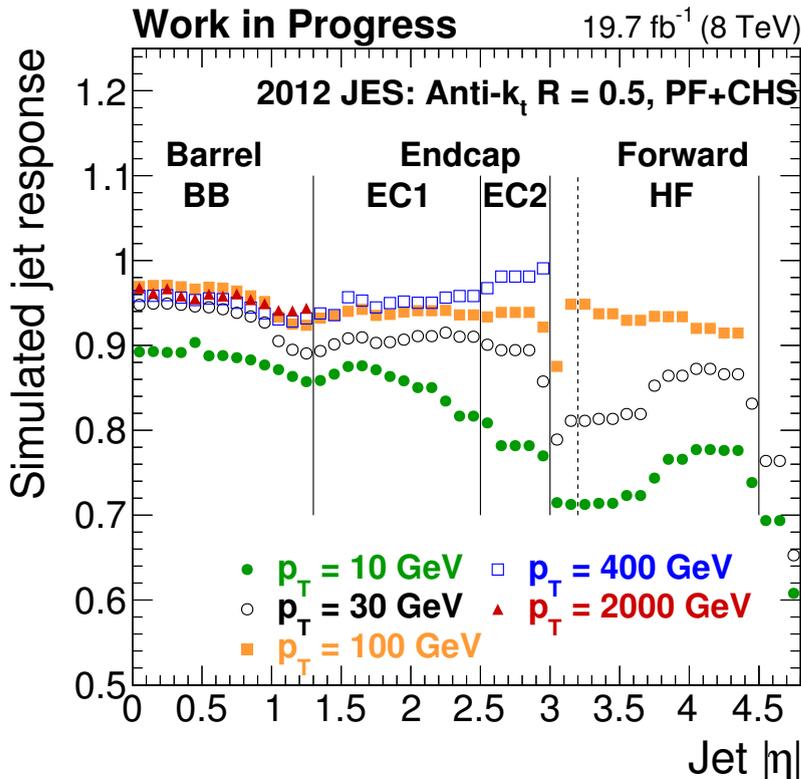
FIGURE 5.4: From Ref. [63], the simulated jet response as a function of pseudorapidity and jet transverse momentum is shown for ak5 PF+CHS jets. The inverse of the response values is the L2L3 correction needed to shift the JES to unity in simulation. The dotted line represents the transition region between HF and endcaps in which parts of the incident jets cannot be reconstructed.

5.4 L1res correction

Taken from Ref. [63], the L1res correction is a modification to the L1 correction, only to be used on data.

There is no guarantee that the pileup modeling in data and simulation or the modeling of the effects of pileup on the detector are accurate. Therefore, a scale factor between data and simulation is estimated from zero bias data and simulation with a random cone method Ref. [64].

In zero bias events, energy depositions from hard interactions are highly improbable. This way, randomly placed cones cluster an average of the offset energy due to pileup in their respective direction. The scale factor is the ratio of the random cone average

pileup offset in data to the same value in simulation. For ak5 PF+CHS jets, the scale factor is shown in Fig. 5.5.
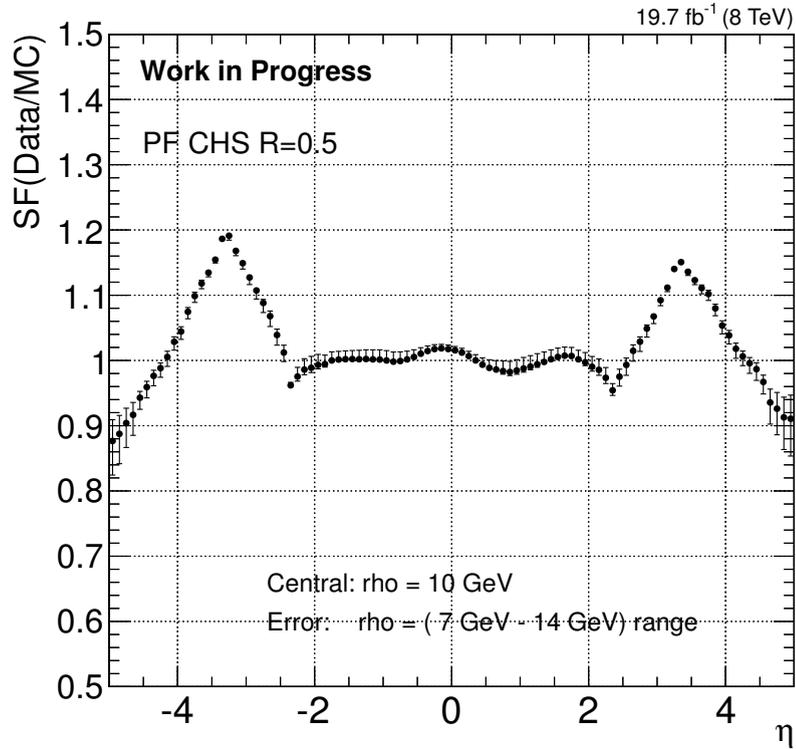
FIGURE 5.5: From Ref. [63], the scale factors between data and simulation derived with the random cone method are shown for ak5 PF+CHS jets. The simulation in the barrel is very accurate, but large scale factors are needed in the endcaps and hadronic forward sections.

Using this scale factor, the L1 correction values are scaled and, for data, this new L1res correction is to be used, instead.

5.5   L2res correction

The L2res correction is used to determine scale factors for the JES between data and simulation relative to a well-understood detector region ($|\eta| < 1.3$). It uses dijets as one of the most common processes at the LHC with a relatively far reach in phasespace. The work is based on the groundwork in Ref. [65] and in 2010 Ref. [64]. It continues the usage of the Kalibri framework (see Ref. [66]) that has been used for the determination of the L2res from 2011 to the present day ($\sqrt{s} = 7$ TeV: Ref. [67]). In this case, the corrections have been derived for $\sqrt{s} = 8$ TeV and are documented over the course of the

run and its reprocessings in: Ref. [68], Ref. [69], Ref. [70], Ref. [71] and finally Ref. [63].

5.5.1   Data and simulated samples used

During the course of the 2012 $\sqrt{s} = 8$ TeV run, several intermediate corrections have been derived and intermediate versions of the data have been used. In this thesis, only the final corrections for the full data with an integrated luminosity of 19.8 fb^{-1} are shown (see Tab. 5.1) with respect to two different cases of simulated samples.

Run	Official CMS Datasets
A	/Jet/Run2012A-22Jan2013-v1/AOD
B	/JetMon/Run2012B-22Jan2013-v1/AOD
B	/JetHT/Run2012B-22Jan2013-v1/AOD
C	/JetMon/Run2012C-22Jan2013-v1/AOD
C	/JetHT/Run2012C-22Jan2013-v1/AOD
D	/JetMon/Run2012D-22Jan2013-v1/AOD
D	/JetHT/Run2012D-22Jan2013-v1/AOD

TABLE 5.1: Reprocessed data samples for the 2012 $\sqrt{s} = 8$ TeV runs. For run B-D, the "JetMon" stream datasets contain low trigger threshold paths (HLT_DiPFJetAve40, 80, 140, 200 and 260) while the "JetHT" stream datasets contain high threshold trigger paths (HLT_DiPFJetAve320 and 400). For run A, the "Jet" stream contains all the aforementioned trigger paths

On the one hand, there are multijet QCD samples with pileup scenario 10 (PU_S10) for two different generators (Pythia 6 and Herwig++). On the other hand, an improvement of out of time pileup description by correctly taking into account three previous bunch crossings to each event has been done for the multijet QCD sample labeled "RD" (run dependent). The simulated datasets are listed in Tab. 5.2.

Process	Official CMS Datasets /QCD_Pt-15to3000_[...]/AODSIM
QCD - Pythia 6	TuneZ2star_Flat_8TeV_pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
QCD - Herwig++	TuneEE3C_Flat_8TeV_herwigpp/Summer12_DR53X-PU_S10_START53_V7A-v1
QCD RD - Pythia 6	TuneZ2star_Flat_8TeV_pythia6/Summer12_DR53X-PU_RD1_START53_V7N-v1

TABLE 5.2: Flat QCD samples without (no annotation) and with run-dependent (RD) corrections to out of time pileup simulation.

5.5.2 Di-jet selection

A range of triggers on di-jet events with varying thresholds (see Tab. 5.3) on $\bar{p}_T$ as defined in Eq. 5.3 is used.

$$\bar{p}_T = \frac{p_T(\text{leading jet}) + p_T(\text{subleading jet})}{2} \tag{5.3}$$

These triggers are labeled *Di* for two, *PF* for a reduced particle flow reconstruction on high level trigger level, *Jet* for jet based and *Ave* for thresholds on $\bar{p}_T$, so DiPFJetAve. Each DiPFJetAve high level trigger is seeded by one level 1 (L1) single jet trigger. Furthermore, the online and offline particle flow reconstruction of the high level trigger differ, depending on the used reconstruction algorithm, the cluster particles used and the cone size used for jet reconstruction. For two different sources of cluster particles and two different jet sizes, three variants and their respective trigger thresholds at which the triggers are 99% effective are compared to the nominal thresholds and shown in Tab. 5.3.

algorithm, jet radius and cluster source	HLT_DiPFJetAve thresholds [GeV]						
	40	80	140	200	260	320	400
ak5PF	60	105	174	242	311	380	468
ak5PF+CHS	62	107	175	242	310	379	467
ak7PF	71	116	190	261	332	401	494

TABLE 5.3: Table of the trigger threshold values in average transverse momentum of the two leading jets. The first row beneath HLT_DiPFJetAve displays the nominal trigger thresholds. The rows beneath contain the trigger threshold per jet radius for anti-kt jets (see Sec. 4.2.1) out of particle flow (PF, see Sec. 4.2.2) clusters with and without charged hadron substraction (CHS, see Sec. 4.2.3). Trigger thresholds have been derived using the approximate method described in Ref. [72] with the lowest threshold being derived by linear extrapolation over all other measured thresholds.

The trigger plateau onset values or thresholds above are also used as bin edges for the L2res correction determination, whenever a $\bar{p}_T$-dependent extrapolation is performed. The trigger plateau fits are shown in Fig. 5.6.

(a) ak5PF jets   (b) ak5PF+CHS jets
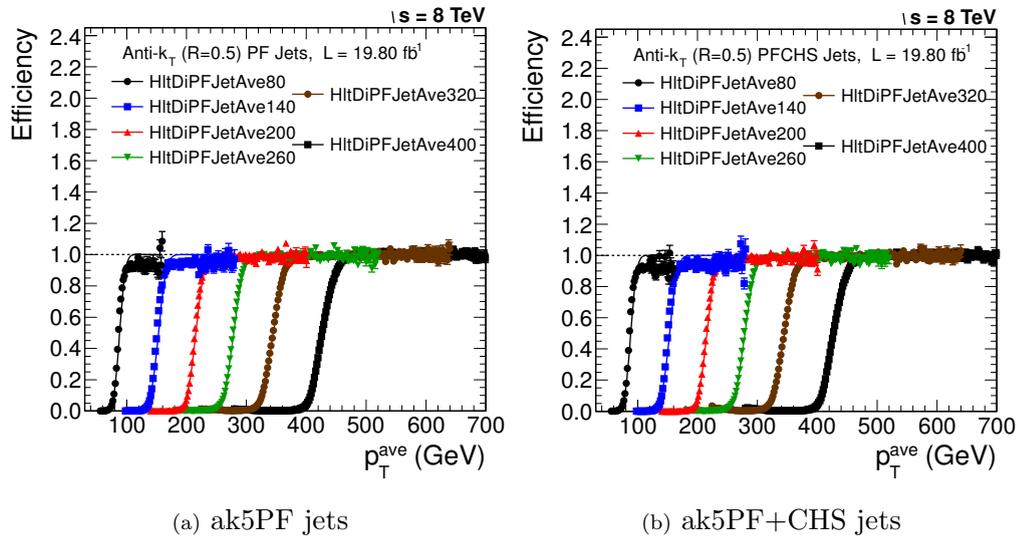
FIGURE 5.6: Trigger efficiencies and fitted plateau values for the dijet selection at $\sqrt{s} = 8$ TeV.

In addition, the bin edges in Tab. 5.4 are also used.

further $\bar{p}_T$ bin edges [GeV]	628	839	1121	1497	2000

TABLE 5.4: Additional loglinearly equidistant bin edges in $\bar{p}_T$ for all algorithms.

The analysis also contains a binning in pseudorapidity. This binning corresponds to the calorimeter tower spacings in the detector and can be further grouped into the barrel, endcaps (EC) and hadronic forwards (HF) regions. The bin edges for the absolute pseudorapidity are listed in Tab. 5.5.

region	$|\eta|$ bin edges						
barrel	0.000	0.087	0.174	0.261	0.348	0.435	0.522
	0.696	0.783	0.957	1.044	1.131	1.218	1.305
EC	1.392	1.479	1.566	1.653	1.740	1.830	1.930
	2.043	2.172	2.322	2.500	2.853	2.964	
HF	3.139	5.191					

TABLE 5.5: Bin edges in $|\eta|$ for the L2residual correction. Segmentations into barrel, endcaps (EC) and hadronic forward (HF) regions are also shown.

Each event being considered as a di-jet event has to pass the following selection requirements:

Trigger: An event has to pass one DiPFJetAve trigger and also pass its determined plateau efficiency thresholds in $\bar{p}_T$ as specified in Tab. 5.3

Two jets: The two leading jets have to pass the loose jet identification criteria (see Sec. 4.2.4)

Barrel jet: At least one of the leading two jets in transverse momentum has to point into the barrel region $|\eta| < 1.3$

Back to back: The two leading jets have to fulfill $\Delta\phi \geq 2.7$, in order to remove QCD multijet events or, to a lesser extent, $Z \rightarrow$ jet+jet events.

Afterwards, eligible events are sorted into a three-dimensional binning that uses pseudo-rapidity, average transverse momentum (see Eq. 5.3) and α, which is defined in Eq. 5.4.

$$\alpha = \frac{p_T(\text{jet } 3)}{\bar{p}_T} \tag{5.4}$$

The variable α is an estimator for the impact of additional radiation, be it either initial state radiation (ISR) or final state radiation (FSR), on the event. The procedure of this three-dimensional binning is:

η Events with one of the two leading jets in transverse momentum pointing into the barrel are sorted into η bins where the bin is determined by the direction of the jet not pointing into the barrel. Furthermore, events with both leading jets pointing into the barrel are considered twice, once in the direction of each jet. The $|\eta|$ bin edges are listed in Tab. 5.5. For the full η binning, symmetric bin edges with negative sign are used, in addition.

$\bar{p}_T$ Each event is sorted into a bin according to its $\bar{p}_T$. The corresponding bin edges are listed in Tab. 5.3 and Tab. 5.4.

α Finally, an inclusive binning in α is performed. Each bin passing, e.g., $\alpha < 0.1$ is considered four times in total, being used for all bin edges of $\alpha < 0.1 \vee 0.2 \vee 0.3 \vee 0.4$

For a value of $\alpha < 0.2$, the resulting selected phasespace that is used in the analysis (each bin at least containing 100 entries) is shown in Fig. 5.7.
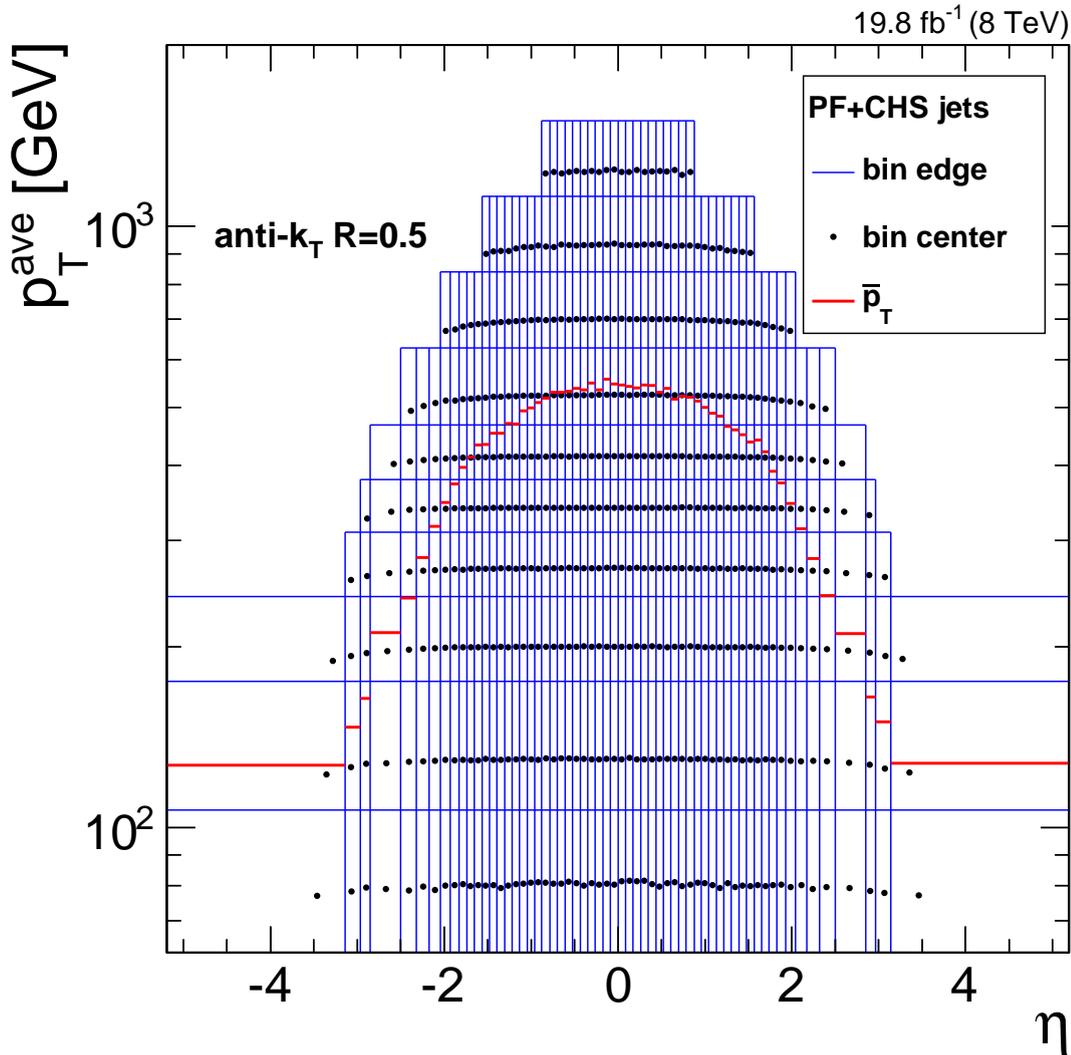
FIGURE 5.7: Diagram of the bin centers in $\bar{p}_T$ and η for di-jet events in bins of the trigger thresholds (see Tab. 5.3) and further logarithmically equidistant bins at high momenta (see Tab. 5.4), as well as the pseuodorapidity bins used for the L2res correction (see Tab. 5.5). Only bins with 100 events or more are shown.

It can be observed that, although the L2res correction is supposed to estimate scaling factors for the JES between simulation and data for the complete detector, the effective range is limited. Three correlated effects cause the effective reach to be lower than desired, especially at far forward regions of the detector. On the one hand, due to the transverse momentum being a projection of the momentum of a jet on the transverse plane, the kinematic limit for the observable transverse momentum is a function of $\cosh(|\eta|)^{-1}$. At high pseudorapidities, this means that a different part of the exponentially falling jet momentum spectrum is observable than in the barrel for a given similar selection requirement on jet transverse momentum. This is the first reason that for similar requirements, different amounts of events per pseudorapidity are selected.

On the other hand, the HLT_DiPFJetAve triggers are heavily prescaled at low jet transverse momenta. Only the lowest jet transverse momenta are observable in the forward regions and they are subject to the same prescales as far more common barrel jets of the same jet transverse momentum range.

The third reason is the requirement on α. As pileup is diffuse in the detector, jets can be clustered out of pileup with low energies or ISR/FSR jets can be enhanced by pileup energy which fluctuates beyond the estimated average being substracted by the L1 correction. This is not problematic, if the hard event is sufficiently energetic. For the forward regions, the geometric necessity of observing jets with low p_T makes these topologies far more vulnerable to failing requirements on α than other η bins with higher $\bar{p}_T$.

The impact of these can be observed in Fig. 5.7 by the bending of the bin centers towards lower $\bar{p}_T$ and the inner edge in η of each bin at high pseudorapidities.

5.5.3 Asymmetry method

A perfect di-jet event consists of two jets back to back in the transverse plane with no additional radiation. In such an event, the transverse momentum of one jet is equal to the other jet.

Given such a situation, any difference in the transverse momenta observed is due to mismeasurements of the true jet transverse momentum. Mismeasurements for each single event cannot be prevented, but using a calibration, the true transverse momentum is measured, on average. For any such calibration, a reference is needed that allows the estimation of an event's precision of measurement. For di-jet events, the problem is that both jets have a jet energy resolution with a nonzero width and, as they are the same objects, there is no ad hoc reference object of good quality. In order to solve this problem, the jet energy scale has to be known for a reference part of the detector. The L3res correction (see Sec. 5.6) is used to determine this absolute scale as a reference in the barrel region ($|\eta| < 1.3$).

A jet pointing into the barrel region is the reference and is to be defined as "tag", a jet pointing anywhere else as "probe". If both jets fall into the barrel region, the event is used twice with the jets alternatingly being defined as "tag" and "probe".

An estimator for the response R and therefore for the jet energy scale as the mean value of the response distribution is the transverse momentum asymmetry or p_T balance defined in Eq. 5.5.

$$R(p_T \text{ balance}) = \frac{1 + \langle \mathcal{A} \rangle}{1 - \langle \mathcal{A} \rangle}; \mathcal{A} = \frac{p_T(\text{probe}) - p_T(\text{tag})}{p_T(\text{probe}) + p_T(\text{tag})} \tag{5.5}$$

The problem with this kind of estimator is that any ISR or FSR disturbance of the jet balance biases the measurement of the JES. This can be corrected for as shown in Sec. 5.5.5 or mitigated by the usage of a better estimator as shown in Sec. 5.5.4.

5.5.4   Missing transverse momentum projection fraction method

The missing transverse momentum projection fraction method (MPF) is an improved estimator of the JES. It is based on the idea that transverse momentum conservation is also observed by ISR and FSR and can be taken into account by including the missing transverse energy (see Sec. 4.6) in a response estimator.

In a perfect di-jet event, the "tag" and "probe" jet as defined in Sec. 5.5.3, have equal transverse momentum, as defined in Eq. 5.6.

$$\vec{p}_T^{\,\text{true}}(\text{tag}) + \vec{p}_T^{\,\text{true}}(\text{probe}) = \vec{0} \tag{5.6}$$

The relation between the measured transverse momentum and true transverse momentum is given by the individual jet response R in the event as defined in Eq. 5.7.

$$\vec{p}_T^{\,\text{meas.}} = \vec{p}_T^{\,\text{true}} \cdot R \tag{5.7}$$

For each event, the vectorial addition of the two measured jet momenta constitutes the missing transverse energy, as described in Eq. 5.8.

$$R(\text{tag}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag}) + R(\text{probe}) \cdot \vec{p}_T^{\,\text{true}}(\text{probe}) = -\vec{\not{E}}_\text{T} \tag{5.8}$$

This equation can be solved in order to get a ratio of the responses $R(\text{probe})/R(\text{tag})$ by first multiplying the overall equation with $\vec{p}_T^{\,\text{true}}(\text{tag})$ as shown in Eq. 5.9.

$$R(\text{tag}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag}) + R(\text{probe}) \cdot \vec{p}_T^{\,\text{true}}(\text{probe}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag}) = -\vec{\not{E}}_\text{T} \cdot \vec{p}_T^{\,\text{true}}(\text{tag}) \tag{5.9}$$

Because the tag jet aligns with itself, $\vec{p}_T^{\,\text{true}}(\text{tag}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag}) = |\vec{p}_T^{\,\text{true}}(\text{tag})|^2 = p_T^{\text{true}}(\text{tag})^2$ is true. Dividing by $p_T^{\text{true}}(\text{tag})^2$ yields Eq. 5.10.

$$R(\text{tag}) + R(\text{probe}) \cdot \frac{\vec{p}_T^{\,\text{true}}(\text{probe}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag})}{p_T^{\text{true}}(\text{tag})^2} = -\frac{\vec{\not{E}}_\text{T} \cdot \vec{p}_T^{\,\text{true}}(\text{tag})}{p_T^{\text{true}}(\text{tag})^2} \tag{5.10}$$

The expression $\frac{\vec{p}_T^{\,\text{true}}(\text{probe}) \cdot \vec{p}_T^{\,\text{true}}(\text{tag})}{p_T^{\text{true}}(\text{tag})^2}$ equals -1, because the tag and probe jet go in opposite directions and the true momenta have equal magnitude. This leads to Eq. 5.11.

$$R(\text{tag}) - R(\text{probe}) = -\frac{\vec{\not{E}}_\text{T} \cdot \vec{p}_T^{\,\text{true}}(\text{tag})}{p_T^{\text{true}}(\text{tag})^2} \tag{5.11}$$

Replacing $\vec{\not{E}}_{\mathrm{T}} \cdot \vec{p}_T^{\,\mathrm{true}}(\mathrm{tag})$ with $\not{E}_{\mathrm{T}} \cdot p_T^{\mathrm{true}}(\mathrm{tag}) \cdot \cos \Delta\phi(\not{E}_{\mathrm{T}},\mathrm{tag})$ and $p_T^{\mathrm{true}}(\mathrm{tag})$ with $p_T^{\mathrm{meas.}}(\mathrm{tag})/R(\mathrm{tag})$ yields Eq. 5.12.

$$R(\mathrm{probe}) = R(\mathrm{tag}) + \frac{\not{E}_{\mathrm{T}} \cdot R(\mathrm{tag}) \cdot \cos \Delta\phi(\not{E}_{\mathrm{T}},\mathrm{tag})}{p_T^{\mathrm{meas.}}(\mathrm{tag})} \tag{5.12}$$

Evidently, this reduces to Eq. 5.13.

$$\frac{R(\mathrm{probe})}{R(\mathrm{tag})} \approx \mathcal{R}(\mathrm{MPF}) = 1 + \frac{\not{E}_{\mathrm{T}} \cdot \cos \Delta\phi(\not{E}_{\mathrm{T}},\mathrm{tag})}{p_T^{\mathrm{meas.}}(\mathrm{tag})} \tag{5.13}$$

There is also an alternative way to formulate the MPF response, as done in Ref. [63]. This alternative definition as shown in Eq. 5.14 and 5.15 is fully equivalent to the MPF definition derived above.

$$\mathcal{B} = \frac{\not{E}_{\mathrm{T}} \cdot \cos \Delta\phi(\not{E}_{\mathrm{T}},\mathrm{tag})}{p_T^{\mathrm{meas.}}(\mathrm{tag}) + p_T^{\mathrm{meas.}}(\mathrm{probe})} \tag{5.14}$$

$$\frac{R(\mathrm{probe})}{R(\mathrm{tag})} \approx \mathcal{R}(\mathrm{MPF}) = \frac{1 + \mathcal{B}}{1 - \mathcal{B}} \tag{5.15}$$

The advantage of this definition is that it uses the same reference variable, namely the average measured momentum of both tag and probe jet, for all purposes, including the MPF response definition. Still, this definition has not been used in this thesis, although the impact of changing the definition has been evaluated and found to be negligible. Fig. 5.8 shows three common cases when trying to measure the JES in di-jet events and explains the advantages of using the MPF method.

(a) mismeasured   (b) FSR   (c) event with diffuse radiation
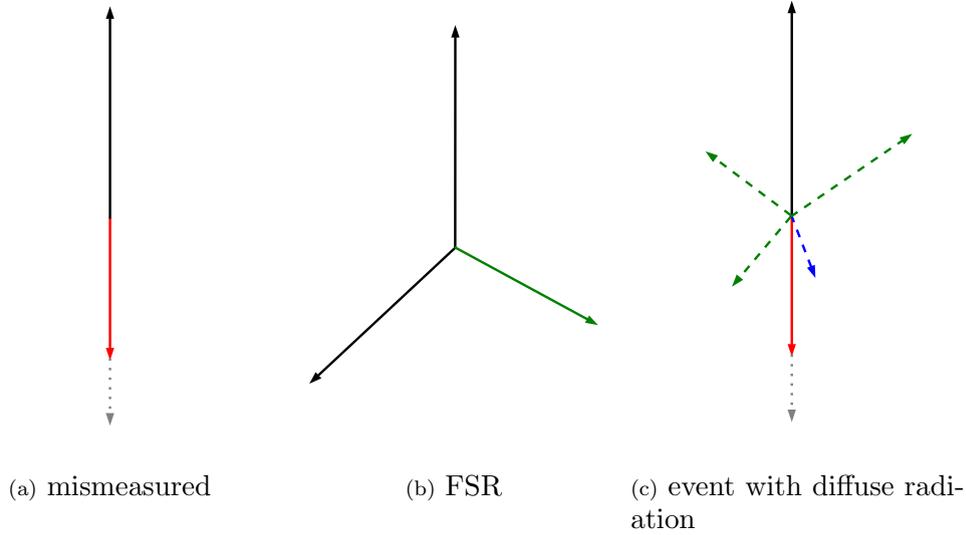
FIGURE 5.8: The three figures show several scenarios common in di-jet events in the transverse plane of the detector. First, Fig. 5.8(a) shows a well-measured jet in black balanced by a mismeasured jet in red. The grey dotted arrow shows the direction and size of the resulting missing transverse momentum. Projecting the transverse momentum on a jet or evaluating the p_T balance yield similar results in this idealized case.

Second, Fig. 5.8(b) shows an event with two well-measured jets in black with one additional well-measured jet due to FSR in green. In such a case, the vectorial sum of all jets is zero. The MPF method will not suffer from any mismeasurement. The p_T balance, on the other hand, will be biased without any additional corrections.

Finally, Fig. 5.8(c) shows an event with a well-measured jet in black balanced by a mismeasured jet in red that is clustered together with a blue jet due to diffuse radiation in the detector. The blue jet itself is balanced by all other diffuse radiation jets in green. In such a complicated but realistic case, the MPF method will still yield the correct result, because it takes into account all of the jets, including diffuse radiation down to threshold induced precision. The p_T balance method, on the other hand, will overestimate the response in such a case, without further corrections.

To summarize, the MPF method is an evolution of the p_T balance, in order to remove or at least mitigate the dependence of the response measurement on additional initial or final state radiation.

5.5.5 Correction for final state radiation

In order to estimate the impact of additional radiation on each event, the variable α is defined in Eq. 5.16.

$$\alpha = \frac{p_T(\text{jet 3})}{\bar{p}_T} \tag{5.16}$$

It is an estimator for additional jet activity in the event, e.g., due to ISR or FSR with the former having a potential and the latter having a definite impact on the jet balance. Pileup and underlying event induced jets also heighten the measured values for α, although they do not necessarily bias the jet balance.

The L2res correction is determined in several bins of $|\eta|$ and $\alpha < x$. For each bin in $|\eta|$, the value of L2res at $\alpha = 0$ is determined by linear extrapolation. The ratio of the correction determined in $\alpha = 0.2$ to the extrapolated value at no additional jet activity is used as a multiplicative correction on the L2res value measured at $\alpha < 0.2$, as shown in Eq. 5.17

$$k_{\text{FSR}} = \frac{L2res(\alpha \to 0)}{L2res(\alpha < 0.2)} \qquad (5.17)$$

There is no theoretical reason for a linear dependence of the correction on α. Furthermore, the statistical correlations of the inclusive bins are not taken into account. Still, the correction factors for the MPF method are sufficiently small (typically $k_{\text{FSR}} < 0.1\%$) that it can be viewed more as a crosscheck on stability than a needed factor. The situation is different for the asymmetry method that has large correction factors.

Examples of extrapolations are shown in Fig. 5.9 for a barrel and an HF region bin. Since

(a) barrel region   (b) HF region
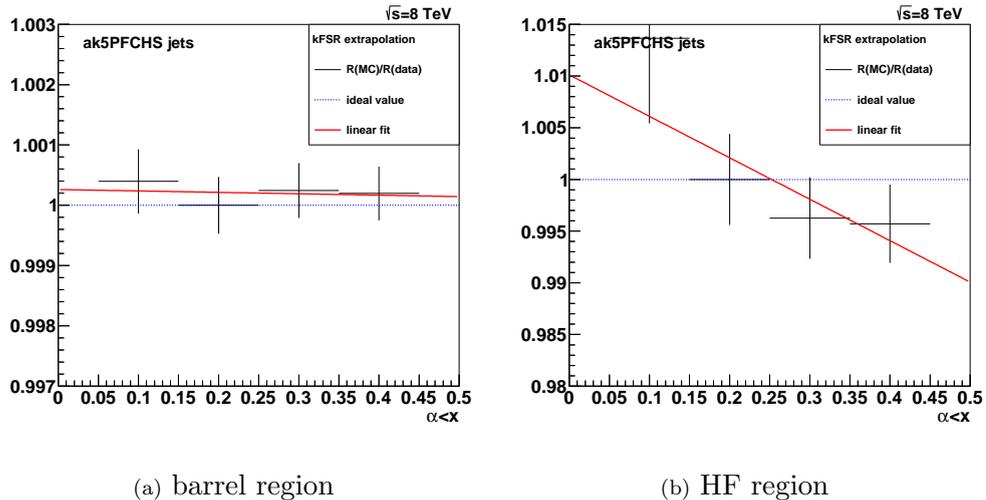
FIGURE 5.9: Measurements of the k_{FSR} for the MPF method. In 5.9(a), an almost flat dependence can be observed. The measured correction factor is smaller than the statistical uncertainty on the bin. For the hadronic forward region in 5.9(b), on the other hand, a substantial $\mathcal{O}(\%)$ correction is determined, but the fit is bad, due to the statistical limitations by the trigger in that region.

the MPF method is stable with respect to final state radiation, one could in principle determine the main correction value at higher cuts on α, in order to reduce the statistical uncertainty on the central correction value. The correction is, in total, already limited by systematical uncertainties in most relevant bins, as will be discussed in Sec. 5.7, so that a further reduction of statistical uncertainties is, albeit possible, not very useful, at

the moment.

All k_{FSR} extrapolations used for the final corrections are listed in App. A.2.

5.5.6 Transverse momentum dependence

After the L1L2L3 or L1resL2L3 corrections, the simulation to data response ratio should be flat with respect to jet transverse momentum, if all previous corrections work as desired and the detector is well understood. This corresponds to a constant fit on a response indicator binned in $\bar{p}_T$.

It is possible that not all detector effects are well understood. In such a case, there would be a dependence of the measured response on $\bar{p}_T$. A loglinear fit defined in Eq. 5.18 is used to estimate such a behaviour.

$$k_{\mathrm{const}} + k_{p_T} \cdot \log(\bar{p}_T) \tag{5.18}$$

The binning used for the fits is shown in Tab. 5.3 and 5.4, in order to avoid mixing events from different trigger paths with different prescales per bin.

At $\sqrt{s} = 7$ TeV, no sizeable loglinear behaviour has been observed. This has changed with respect to $\sqrt{s} = 8$ TeV. All corrections discussed in this thesis use loglinear fits and therefore have $\bar{p}_T$-dependent correction factors for all bins except for the region $|\eta| > 2.964$, where the numbers of events are insufficient to have a reliable fit result. Still, the loglinear fits are used even there for the relative p_T uncertainty discussed in Sec. 5.5.7.

Examples of loglinear fits are shown in Fig. 5.10 for a barrel and an endcap region bin.

(a) barrel region

(b) endcap region
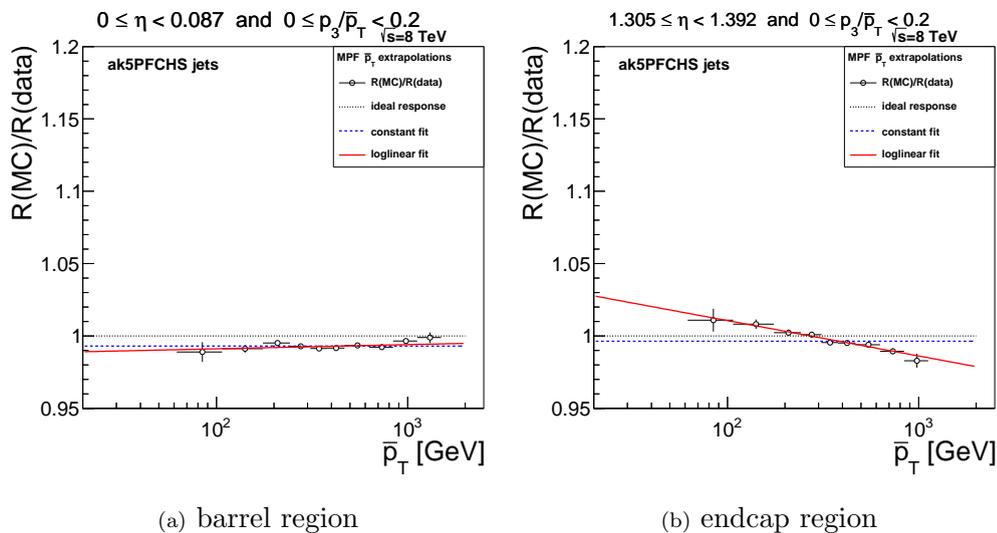
FIGURE 5.10: In 5.10(a), an almost flat dependence can be observed. This is different in the endcaps, where sizeable $\bar{p}_T$ dependencies are observed, like in 5.10(b).

All $\overline{p}_T$ extrapolations used for the final corrections are listed in App. A.3.

5.5.7 Uncertainty sources

Several uncertainties are taken into account for the L2res correction. As described in Ref. [63], soft radiation, jet energy resolution, the relative $\overline{p}_T$ dependence, statistical effects and time dependencies are taken into account. The aforementioned descriptions are summarized, below.

The soft radiation uncertainty is estimated by comparing Herwig++ and Pythia6 for the true simulated response. A level of closure within 0.5% is observed for $|\eta| < 3$ and of 1.5% for $|\eta| > 3$ and assigned as a systematic uncertainty. It is labeled *RelativeFSR*. The JER differences between data and simulation are measured as described in Ref. [73] and then applied on simulated events to scale the jet energy resolution to similar values as in data. Relative differences in the JER between the tag and the probe jet propagate to the measurement of the JES. Therefore, the uncertainty is evaluated by over- and underscaling all simulated jets within the uncertainties of the scaling factors and then taking the difference in the resulting correction factors per bin in η. It is labeled *RelativeJER*.

Because the reason for the $\overline{p}_T$ dependence of the L2res correction is unknown, half the difference between constant and loglinear correction sets is taken as a systematic uncertainty. This is typically the dominant contribution to the uncertainty on the L2res correction and could be substantially reduced or even dropped by an understanding of the origin of the observed $\overline{p}_T$ dependence. It is labeled *RelativePt*.

The statistical uncertainty per bin is mostly negligible for the barrel region and small for the endcaps. It is relevant for the hadronic forward region, where trigger thresholds and the detector geometry lead to a shortage of usable events. Therefore, the events of both detector regions in $|\eta|$ are used, together, in order to reduce the statistical uncertainty in this one region of $|\eta| > 2.964$. Still, the statistical uncertainty in that region can be as large as 2.5%. This uncertainty source is labeled *RelativeStat*.

Radiation damage in the endcaps suffered by the ECAL and HCAL during data taking may have caused a time dependence for the L2res correction. It is only certain that the energy scale of the calorimeters changed over time, although other causes like recalibrations cannot be ruled out. As no time dependence is propagated to the correction factors, uncertainties are assigned as root mean square (RMS) of the η dependent correction factors for a set of ten run ranges. It is labeled *TimeEta*.

The uncertainties for JER, $\overline{p}_T$ and number of events per bin are assumed to be correlated within the following regions:

barrel $0 \leq |\eta| < 1.3$

 EC1 $1.3 \leq |\eta| < 2.5$

 EC2 $2.5 \leq |\eta| < 3$

  HF $3 \leq |\eta| < 5.2$

Uncertainties on time, jet flavor and soft radiation are considered to be fully correlated in η.

5.6   L3res correction

The L3res correction as described in Ref. [63] determines the absolute simulation to data ratio of the JES in the barrel region of $|\eta| < 1.3$. It is determined in several channels with different typically well-reconstructed objects as reference: $Z \to \mu^+ \mu^-$+jet, $Z \to e^+ e^-$+jet, γ+jet and multijet balance. To the L3res, the same methods as for the L2res like MPF, p_T-balance and the application of a k_{FSR} do apply. Only the multijet channel is special in the sense that it's events are used twice: Two jets are used to balance a third jet that is more highly energetic. Depending on whether the leading jet is used as a reference or the two subleading jets are used, very high or intermediate transverse momenta can be probed.

Each channel separately determines the absolute scale in the barrel region as a function of p_T of the reference object and applies a correction for final state radiation on that result. All measurements are then subject to a global fit in which all contributions are varied within their respective unertainties, as shown in Fig. 5.11.
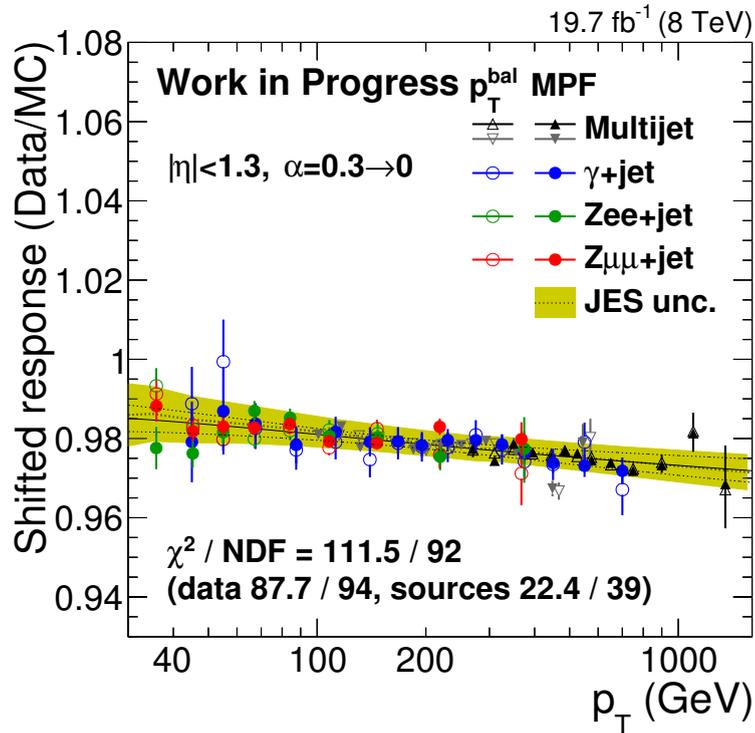
FIGURE 5.11: Taken from Ref. [63], the global fit of the L3res in all channels is shown. A different kinematic reach can be observed for the different channels.

A small deviation from unity with a loglinear p_T dependence is observed.

5.7 Results for the L2res corrections at $\sqrt{s} = 8$ TeV

In the following, the results of the final set of $\sqrt{s} = 8$ TeV corrections are shown for the L2res correction. All correction values are listed in App. A.1, as well as the equation needed to apply those values to jets in data (see Eq. A.1).

5.7.1 Independence of FSR modeling

As described in Sec. 5.5.7, the systematic uncertainty due to soft radiation is estimated by comparing sets of L2res corrections relative to Herwig++ and Pythia6 simulated samples. To estimate the impact of soft radiation, the k_{FSR} is used. In Fig. 5.12(a) a large difference in the modeling can be seen by the different correction factors for p_T balance. Despite this sizeable difference, the MPF method is not affected and all combinations of methods and generators do agree to high precision, as shown in Fig. 5.12(b).

(a) k_{FSR} values   (b) L2res values
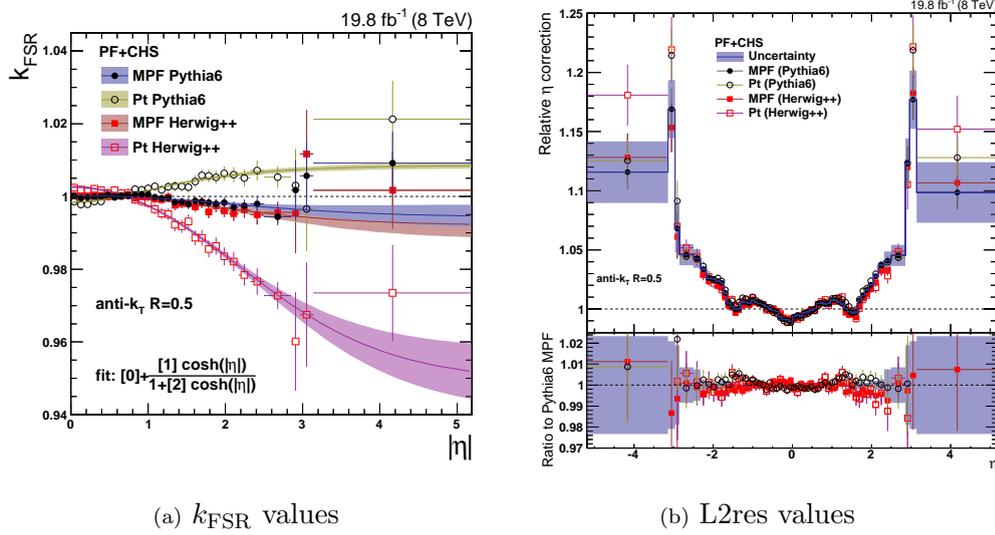
FIGURE 5.12: The final state radiation correction factors are shown in 5.12(a) and the final corrections in 5.12(b) for Pythia6 and Herwig++ determined by the p_T balance or MPF method. The reference simulation samples do not have run dependent corrections for this study, because to date no Herwig++ sample with these corrections has been generated.

A fit of the final state radiation correction is shown that takes into account the comparison of two objects following a kinematic limit of $\cosh(|\eta|)^{-1}$. Differences of up to 5% for the asymmetry method for the k_{FSR} are opposed to differences of up to 0.25% for the MPF method. Half of the difference of the fits for the MPF method between Pythia6 and Herwig++ is used to determine the *RelativeFSR* uncertainty.

In general, Fig. 5.12 demonstrates that the machinery to correct for final state radiation works to high precision and that the application of the MPF method renders the same machinery dispensable.

5.7.2   Transverse momentum dependence impact

Unlike observations in previous years, the $\bar{p}_T$ dependence for the $\sqrt{s} = 8$ TeV L2res corrections is sizeable. For the first time, it has been propagated to the actual corrections. The size of the $\bar{p}_T$ dependence per bin of η is shown in Fig. 5.13.
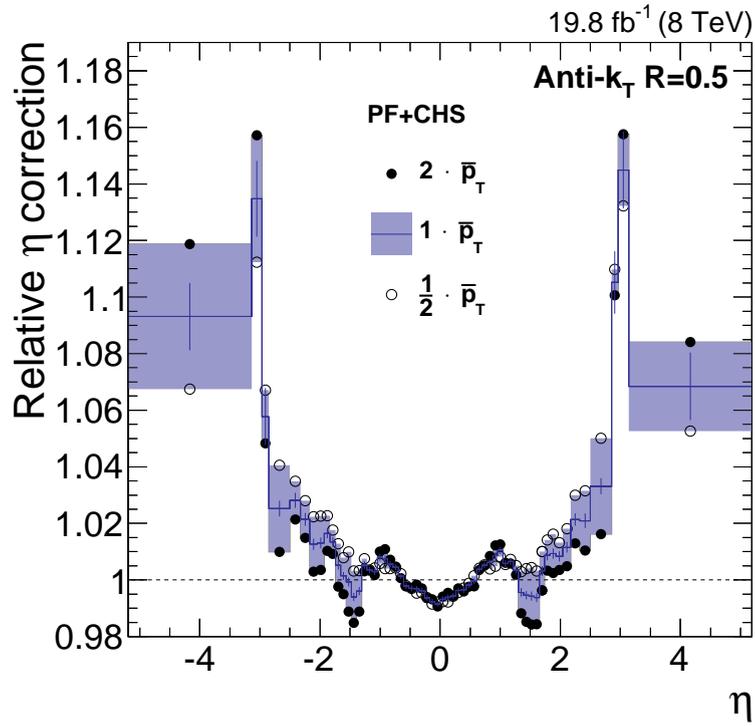
FIGURE 5.13: The L2res correction factors are shown per bin in η for the average, twice and half the $\bar{p}_T$ value. Sizeable dependencies can be observed for the endcaps and the HF region.

5.7.3 Impact of type I corrected MET

For the L2res correction, only a limited type I corrected MET could be applied for simulation (see Eq. 4.4) and data (see Eq. 5.19), respectively.

$$\not{E}_\mathrm{T} = \not{E}_\mathrm{T}^{\mathrm{raw}} + \sum_i \vec{p}_T(\mathrm{jet}_i) \cdot (L1res \cdot L2L3 - L1res) \tag{5.19}$$

This is caused by the circumstance that an application of the L2L3res to the MET would necessitate several iterations of deriving the L2res correction, if the impact of the type I correction on the MET was large. Aside from such a procedure being impractical, it would be unnecessary, if the impact of the corrections on MET would be sufficiently small.

In order to test the impact, the correction values have been derived with and without type I corrections, respectively. A comparison of both values is shown in Fig. 5.14.
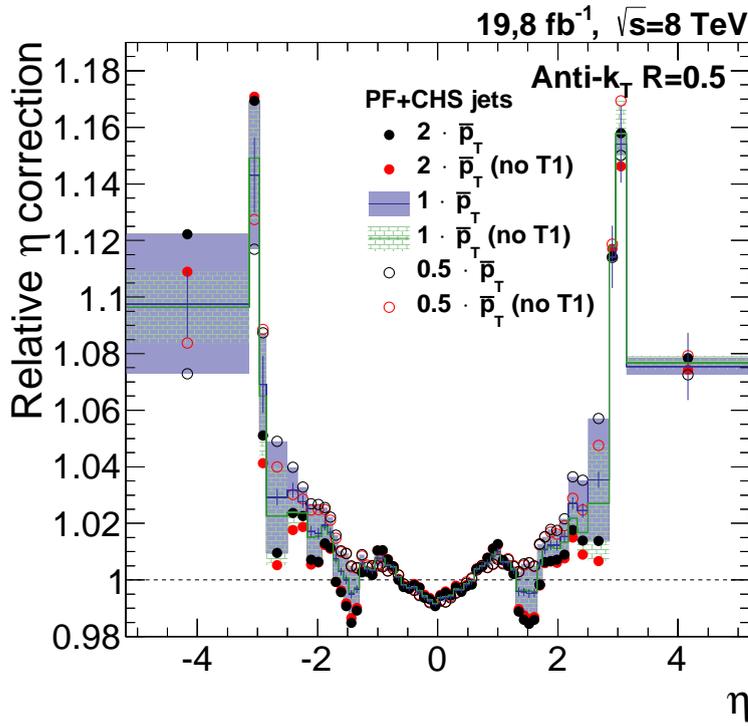
FIGURE 5.14: Comparison of determining the L2res correction with respect to type I MET and raw MET (no T1). Values are shown depending on $\bar{p}_T$. Small deviations are observed in the endcap region, but the overall impact is limited.

It can be shown that few bins are actually affected by correcting the jets for the MET calculation or not. All of these bins are in the endcap region so that the impact is limited to roughly $2.1 < |\eta| < 3$.

To summarize, it is unknown what the precise impact of the missing cycle of iterations towards a more or less perfect corrections is, but the region in which it would matter is known as well as the expected order of magnitude.

A dedicated systematic uncertainty has not been assigned for this effect at the current time.

5.7.4 η asymmetry

In previous years, the L2res correction has been derived in bins of $|\eta|$. For the $\sqrt{s} = 8$ TeV corrections, it has been shown that some bins in the endcap region exhibit a significantly asymmetric behaviour. That behaviour with bins outside a 96.2% confidence level range marked in red is shown in Fig. 5.15.
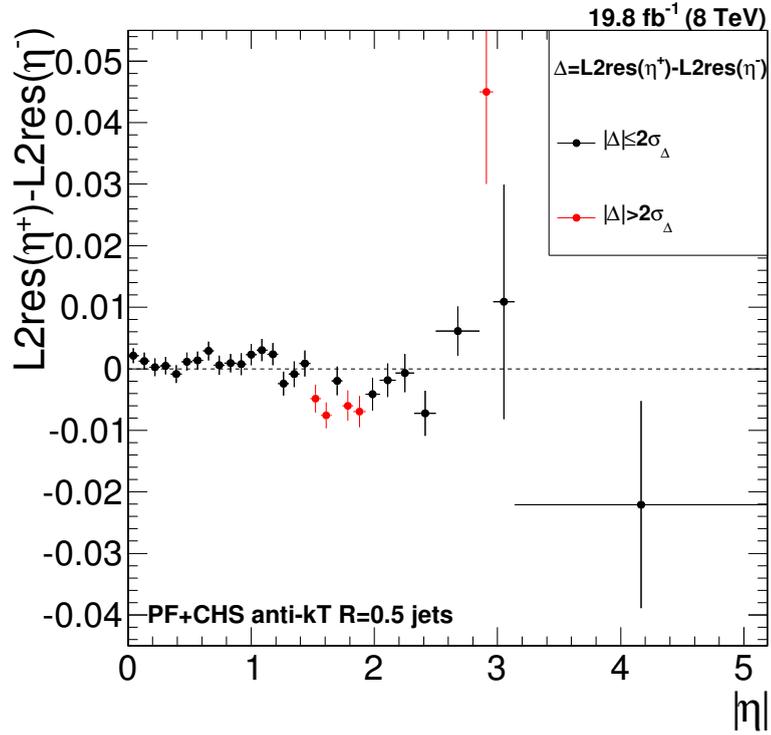
FIGURE 5.15: The difference of the L2res correction values at average $\bar{p}_T$ per bin is shown for the positive and the negative detector hemispheres. A sizeable difference is observed in the region of $1.5 < |\eta| < 2$ region.

For 30 bins, about one correction value would be expected to be in a two σ range due to the number of events. Instead, five bins with such a behaviour are observed. Therefore, it has been decided to make full use of the number of events availabe and determine the corrections in η with the exception of the two bins at $2.964 \leq |\eta| < 5.191$ where the number of events is insufficient.

5.7.5   Stability over time and uncertainty reduction

One of the previously sizeable uncertainties has been due to the measurement of the jet energy resolution in data. In Fig. 5.16(a), the current JER scaling values are shown and in Fig. 5.16(b) the impact on the measurement of the L2res.

(a) JER factors from Ref. [73]

(b) JER variation on L2res
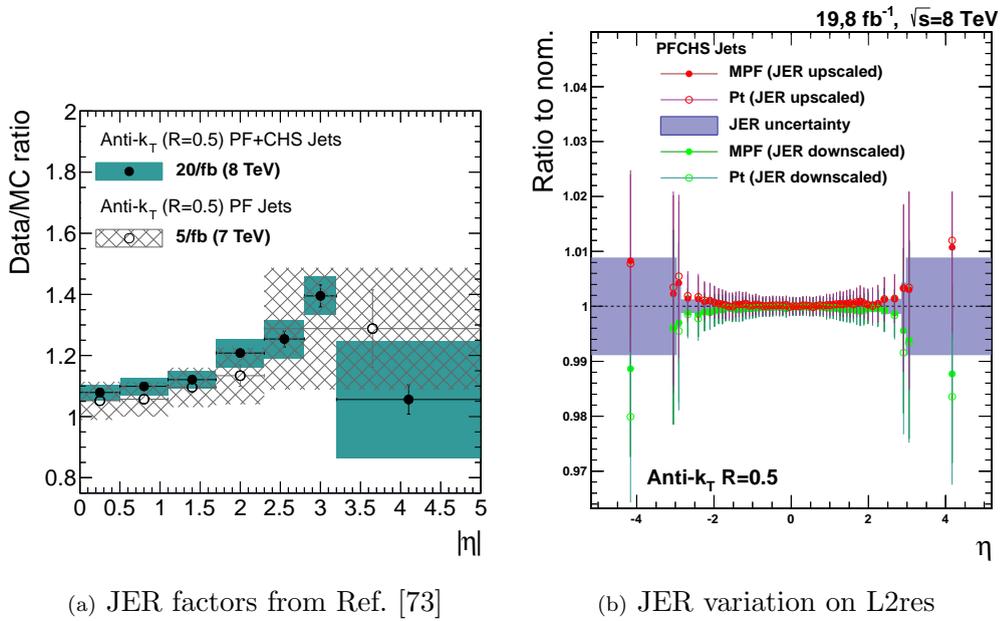
FIGURE 5.16: Fig. 5.16(a) shows the improvement of the determination of smearing factors from $\sqrt{s} = 7$ TeV to 8 TeV. The reduced uncertainties on the JER directly propagate to the resulting JER uncertainties in the L2res shown in Fig. 5.16(b).

The uncertainties stemming from JER variations are now negligible anywhere but in the very forward regions.

Furthermore, the comparison to the $\sqrt{s} = 7$ TeV run can be sought, as shown in 5.17.

(a) low p_T L2L3res

(b) intermediate p_T L2L3res
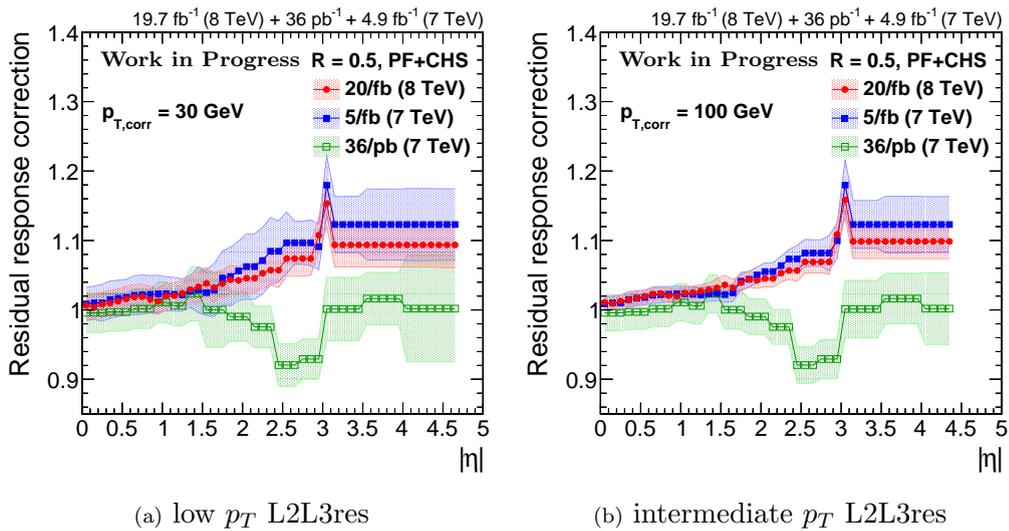
FIGURE 5.17: Taken from Ref. [63], the L2L3res is shown for different years and center of mass energies. It is unknown why the first run at $\sqrt{s} = 7$ TeV with 36 pb^{-1} has such different corrections. That fact aside, the overall corrections have been stable at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV and uncertainties could be reduced, especially at low transverse momenta.

In comparison to previous iterations, the larger number of events due to higher integrated luminosity and the continuous usage of the MPF method have reduced uncertainties, especially at lower transverse momenta. Still, the uncertainty of the measurement in the endcaps and barrel is completely dominated by the uncertainty due to the p_T dependence of the measured L2res correction for unknown reasons, as shown in Fig. 5.18.

(a) low p_T L2res   (b) outer endcap L2res
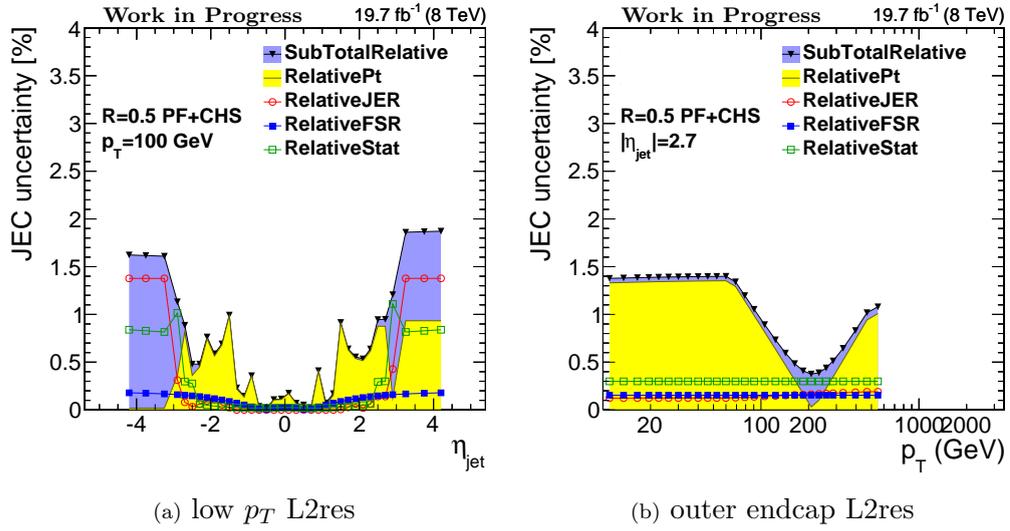
FIGURE 5.18: Taken from Ref. [63], the L2res uncertainties are shown for low transverse momenta and a bin in the outer endcaps. Uncertainties in the barrel and endcaps are completely dominated by the *RelativePt* uncertainty. The number of events gets critically low in the HF region, causing several uncertainty sources to rise, substantially.

If one determined the reason for the $\bar{p}_T$ dependence, used that function for the $\bar{p}_T$ extrapolations and then replaced this *RelativePt* uncertainty with, e.g., fit quality arguments, the uncertainty due to the L2res would be essentially zero in $|\eta| < 3$.

For the HF, the situation is different. Here, the statistical limitations cause huge uncertainties and also the JER has a sizeable impact. In order to fix this misbehaviour, a dedicated set of forward plus barrel di-jet triggers with low prescales would be one way. Another improvement could be reached by extending the well-defined response in the barrel region to the endcaps with a first calibration, then using this new extended reference region for a second determination of the L2res correction. The corrections derived from the additional events with the endcaps as a reference region could then be combined with the corrections derived from events with the barrel as a reference region to improve the precision in the very forward detector regions.

Overall, the measurement of the relative residual JES corrections for $\sqrt{s} = 8$ TeV has been performed to high precision. Still, the superior statistics expected in the next run of the LHC gives rise to further possible applications in the future. It would be possible

to further bin the L2res correction in ϕ as well as η and $\bar{p}_T$, enabling a calibration of all detector elements to the level of the detector granularity.

Chapter 6

Jets misreconstructed as τ_h in a VBF $\tau_\mathrm{h}^\pm\tau_\mathrm{h}^\pm$ SUSY search

As jets are the most common physical objects at the LHC, even small probabilities for jets being reconstructed as a different object can translate into large backgrounds in searches for new physics. In this chapter, the properties of jets faking hadronically decaying τ leptons (τ_h) are investigated and simulation-based methods are derived to predict the shape of the contribution of topologies with at least four jets in a like sign di-τ_h vector boson fusion (VBF) search for supersymmetry.

At the end of the chapter, a yet untested partially data-driven setup to improve the demonstrated method is outlined.

6.1 Search for supersymmetry in like sign di-τ_h vector boson fusion final states

Supersymmetry (SUSY) has been searched for in various channels and the leftover parameter space for natural SUSY, at least, is dwindling. One part of the parameter space of, e.g., the minimal supersymmetric model (MSSM) that is not yet thoroughly explored is the particularly challenging area of compressed spectra. In these scenarios, the mass spread between sparticles is very small. Any standard model particle in a decay chain will be affected by the boost of the decaying particle and the energy available by the mass difference of the incoming particle to the outgoing particles. If the mass differences are small, the resulting outgoing particles will be soft, unless they are boosted to high momenta. This affects reconstructible objects as well as not reconstructed objects like neutrinos, leading to the amount of $\not{E}_\mathrm{T}$ in compressed spectra scenarios being dominated

by the lightest stable supersymmetric particle mass. Soft jets cannot be reliably triggered on, as they are drowned in soft QCD processes. The same reasoning applies to hadronically decaying τ leptons and, to some extent, to electrons. Only with muons, one can go down to the necessary low thresholds on the transverse momentum.

A way to circumvent this problem is to use a different production mode. Typically, in vector boson fusion (VBF), two hard jets with a large rapidity gap and two possibly soft leptons are produced (see Fig. 6.1). In theory, this allows triggering on and tagging of the hard jets in the event. Unfortunately, for the 8 TeV run, no appropriate trigger is available. The CMS detector trigger logic for jets differentiates between $|\eta| < 3$ and $|\eta| \geq 3$ for clustering jets. A considerable part of the expected signal's forward jets would need to be reconstructed and triggered in the region around $|\eta| \approx 3$. As not all jet triggers do consider that region, this is already problematic. Even more problematic are the high prescales for jet triggers. Although there are single jet triggers available that could be used to trigger on one jet, the prescales are so huge that a search with these triggers is not feasible. Lastly, a vector boson fusion jet trigger was implemented for parked data sets, but could be not be used, since it triggered on a level 1 $\not{E}_T$ component that was fully efficient at roughly $\not{E}_T \geq 250$ GeV. That value would have been far beyond the expectation for the signal models considered.

Still, the first searches in this channel are performed in a multitude of τ decay final states: $\mu\mu$, μe, $\mu\tau_h$, $\tau_h\tau_h$ (τ_h are hadronically decaying τ) in opposite and like sign variants. All these analyses share common requirements on the two forward jets and require two central leptons. Differences arise due to trigger thresholds being significantly more benign for the analyses triggering on μ, as well as due to different background compositions that motivate various requirements on $\not{E}_T$. The internal documentation for the di-τ_h channels can be found in Ref. [74] and has been written in collaboration with D. Marconi for the like sign di-τ_h case.

6.1.1 Central and vector boson fusion selections

There are two different parts of the event to apply selections on. On the one hand, the vector boson fusion (VBF) part is selected, consisting of two jets with a high rapidity gap, and on the other hand, the central part of two like sign hadronically decaying τ leptons is selected. This can be seen in the two leading order Feynman diagrams for the VBF production channel in Fig. 6.1, where two jets and two leptons are produced.

(a) up-type VBF production   (b) down-type VBF production
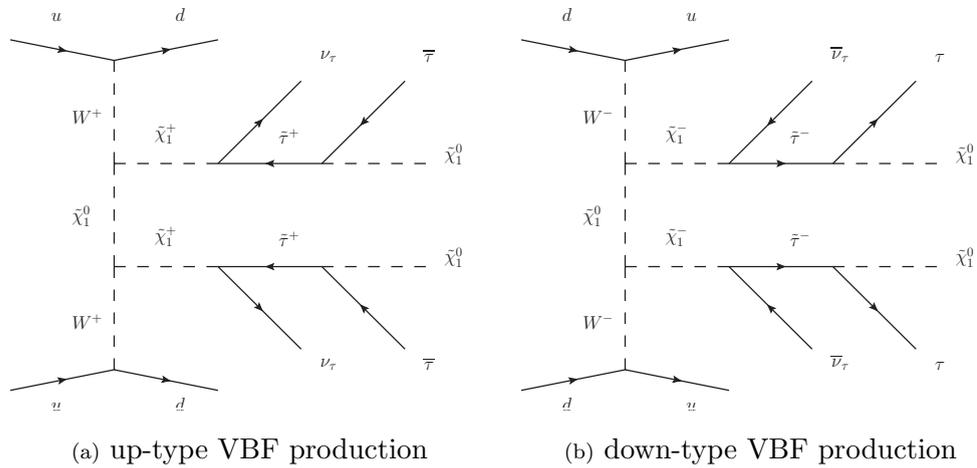
FIGURE 6.1: Feynman diagrams of leading order production processes for the like sign di-τ_h channel.

The central selections for all events are:

- The HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Prong1_v* trigger

- Exactly 2 τ_h of a given isolation (see Sec. 4.5 for definitions) with $p_T^{\tau_h} \geq 45$ GeV

- $\not{E}_T \geq 30$ GeV

- At least 2 jets not within $\Delta R(\text{jet}, \tau_h) \leq 0.3$, with $p_T^{\text{jet}} \geq 30$ and within $|\eta| \leq 5.0$

- No jet with a b tag (see Sec. 4.2.5 for definitions) not within $\Delta R(\text{jet}, \tau_h) \leq 0.3$

The high level trigger choice is unfortunate, but without an alternative for the $\sqrt{s} = 8$ TeV run as it is the only available trigger without a prescale that has a high signal acceptance, but also the least competitive trigger in comparison to other lepton triggers. The requirement on τ_h transverse momentum is chosen, in order to have a fully efficient trigger. As indirectly indicated, it is the biggest limitation for the sensitivity of the analysis with respect to compressed spectra exhibiting soft leptons at the current time. The $\not{E}_T$ is produced by the τ neutrinos of the $\tilde{\tau}$ production and the two decaying τ_h, and for SUSY events by the lightest stable supersymmetric particle (LSP), in this case assumed to be the $\tilde{\chi}_1^0$. This can be used to set apart events of interest from QCD, as the latter contains a negligible amount of $\not{E}_T$ by neutrinos produced during hadronization of jets or a possibly major amount of $\not{E}_T$ produced by mismeasurements of the jets. Moreover, exactly two τ_h are required, as a sizeable number of signal events with three or more τ_h is not expected and all cases with three or more τ_h are neglected for the background estimation.

In the VBF selection, any combination of two jets not matched to τ_h need to fulfill:

- $\Delta\eta(\text{jet, jet}) \geq 4.2$

- $M^{\text{jet, jet}} \geq 250$ GeV

- $\text{sign}(\eta^{\text{jet 1}} \cdot \eta^{\text{jet 2}}) = -1$

The invariant mass of the dijet-system $M^{\text{jet jet}}$ is a cumulative property of the two jet four-vectors. Thus, this requirement is strongly correlated with the rapidity gap requirement of $\Delta\eta(\text{jet, jet})$.

The central and VBF selections are kept separate, as the VBF selection is inverted for some control regions as part of the estimation for QCD background, described in the Sec. 6.1.3.

6.1.2 Used SM and signal samples

Lists of samples and trigger requirements, taken from Ref. [74]:

Process	Official CMS Datasets /DY*/AODSIM
$Z \to \tau\tau$	ToTauTau_M-20_CT10_TuneZ2star_v2_8TeV-powheg-tauola-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v2
$Z \to \mu\mu$	ToMuMu_M-20_CT10_TuneZ2star_v2_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ee$	ToEE_M-20_CT10_TuneZ2star_v2_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll(10 < m_{ll} < 50)$	JetsToLL_M-10To50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll(m_{ll} > 50)$	JetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll + 1jets$	1JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll + 2jets$	2JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll + 3jets$	3JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll + 4jets$	4JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$Z \to ll$ EWK	JJ01JetsToLL_M-50_MJJ-200_TuneZ2Star_8TeV-madgraph_tauola/Summer12_DR53X-PU_S10_START53_V7A-v1

TABLE 6.1: Drell Yang simulated samples.

Process	Official CMS Datasets /W*/AODSIM
$W + 0$ jets	JetsToLNu_TuneZ2Star_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v2
$W + 1$ jet	1JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$W + 2$ jets	2JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$W + 3$ jets	3JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
$W + 4$ jets	4JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1

TABLE 6.2: W boson plus additional jets simulated samples.

Process	Official CMS Datasets */TTJets*/AODSIM*
$t\bar{t}$	_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7C-v1

TABLE 6.3: Standard model top production simulated sample.

Process	Official CMS Datasets */AODSIM*
$WW(\to 2l2\nu)$	WJetTo2L2Nu_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7C-v1
W^+W^+	/WpWpqq_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
W^-W^-	/WmWmqq_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1
WW double scattering	/WW_DoubleScattering_8TeV-pythia8/Summer12_DR53X-PU_S10_START53_V7A-v1
WW EWK	/WWjjTo2L2Nu_8TeV_madgraph_qed6_qcd0/Summer12_DR53X-PU_S10_START53_V19-v1
$WZ(\to 2q2\nu)$	/WZJetsTo2Q2Nu_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1
$WZ(\to 2l2\nu)$	/WZJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1
$WZ(\to 3l)$	/WZJetsTo3L_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ(\to 2q2\nu)$	/ZZJetsTo2Q2Nu_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ(\to 2l2\nu)$	/ZZJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ(\to 2l2q)$	/ZZJetsTo2L2Q_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1
$ZZ(\to 4l)$	/ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1

TABLE 6.4: Standard model production of two vector bosons simulated samples.

Process	Official CMS Datasets */VBF_*/AODSIM*
$H \to WW(\to 2l)$	HToWWTo2LAndTau2Nu_M-125_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$H \to ZZ(\to 2l2\nu)$	HToZZTo2L2Nu_M-120_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$H \to ZZ(\to 2l2q)$	HToZZTo2L2Q_M-125_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$H \to ZZ(\to 4l)$	HToZZTo4L_M-125_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$H \to ZZ(\to 4\nu)$	HToZZTo4Nu_M-120_8TeV-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1
$H \to \tau\tau$	HToTauTau_M-125_8TeV-powheg-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1

TABLE 6.5: Standard model Higgs production by vector boson fusion simulated samples.

Process	Official CMS datasets */AODSIM
$bg \to tW^-$	/T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
$\overline{b}g \to \overline{t}W^+$	/Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
$q'b \to qt$	/T_t-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
$q\overline{b} \to q'\overline{t}$	/Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
$q\overline{q}' \to t\overline{b}$	/T_s-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1
$\overline{q}q' \to \overline{t}b$	/Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1

TABLE 6.6: Single top simulated samples.

```
CMSSW collection label:  TriggerResults_HLT
CMSSW type:  edm::TriggerResults
```

HLT_DoubleMediumIsoPFTau35_Trk5_eta2p1_Prong1_v2	OR
HLT_DoubleMediumIsoPFTau35_Trk5_eta2p1_Prong1_v3	OR
HLT_DoubleMediumIsoPFTau35_Trk5_eta2p1_Prong1_v4	OR
HLT_DoubleMediumIsoPFTau35_Trk5_eta2p1_Prong1_v6	OR
HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Prong1_v1	OR
HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Prong1_v3	OR
HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Prong1_v4	

TABLE 6.7: Explicit list of trigger paths required.

Taken from Ref. [75], private signal samples have been generated by using Madgraph to generate LHE files. The LHE files contain lists of events and particles produced therein. FastSim CMS software has been used to produce AODSIM events.

For all datasets, a common Physics Analysis Toolkit (PAT) [76] sequence has been used to generate a PAT format, then further reduced in size by the ntuple producer [77].

6.1.3   Standard Model background estimation

Four types of background are considered. Firstly, the irreducible standard model vector boson fusion backgrounds producing two jets and two τ_h, shown in Fig. 6.2.

(a) up-type VBF SM production   (b) down-type VBF SM production
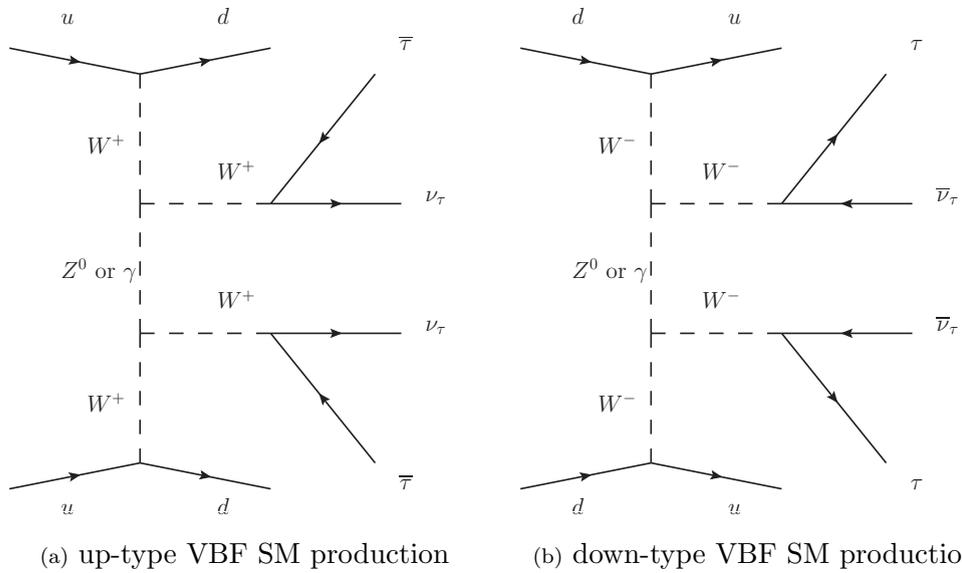
FIGURE 6.2: Feynman diagrams of irreducible backgrounds for the vector boson fusion like sign di-τ_h supersymmetry search.

The number of events contributed by this irreducible background is minor, as the production cross sections for these processes is very small. Furthermore, these processes are expected to be well-modeled by the simulation as they are purely electroweak leading to the expectation that such backgrounds can be reliably substracted. Summarily, they are of minor importance and can be taken directly from simulation.

Secondly, there are charge misreconstructed standard model backgrounds, shown in Fig. 6.3.

(a) VBF Z^0 SM production by Z-bosons (b) VBF Z^0 SM production by W-bosons

(c) VBF H SM production by Z-bosons (d) VBF H SM production by W-bosons

(e) Di-leptonic SM decay of top-quark pairs: Both b jets have to fail the b tag and one of the τ needs to have charge misreconstruction to enter the selection
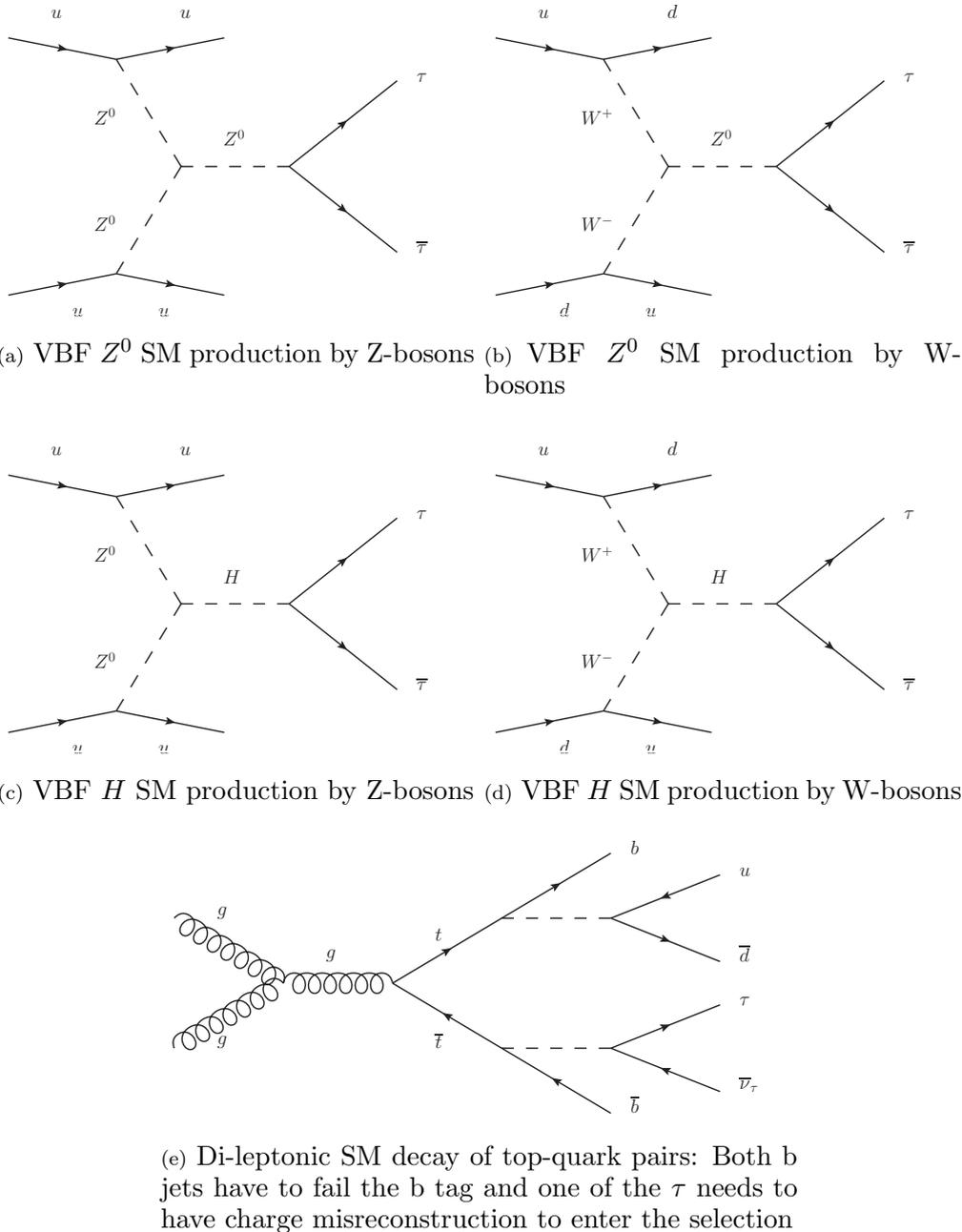
FIGURE 6.3: Feynman diagrams of standard model backgrounds for the vector boson fusion like sign di-τ_h supersymmetry search, in which one τ_h with a charge misidentification is sufficient to pass the like sign selection criteria.

The charge misreconstruction rate of τ_h leptons is not well known, but can be estimated in $M^{\tau\tau}$ distributions, as most of these contributions would stem from Z $\to \tau_h\tau_h$ processes. Thus, the invariant mass distribution of the charge misreconstructed di-τ_h-system would be close to the Z-boson peak and clearly visible in any like sign control region, where a QCD purity of $\approx 90\%$ is expected. On the other hand, the γZ interference continuum

is expected to not contribute significantly, because the expected event rate is so low that even if any events are observed, the chance of any event on the Z-boson peak to be reconstructed is highest.

A notable exception is the di-leptonic decay of top quark pairs that would not show up in an invariant mass distribution. However, this can be suppressed by requiring no b tag on any jet not matched to τ_h in the event. It will be shown in Sec. 6.3 that any such contribution appears to be minor and negligible.

Thirdly, standard model VBF processes with three leptons can contribute, where one lepton is too soft and falls below the reconstruction thresholds, examples shown in Fig. 6.4.

(a) up-type   (b) down-type
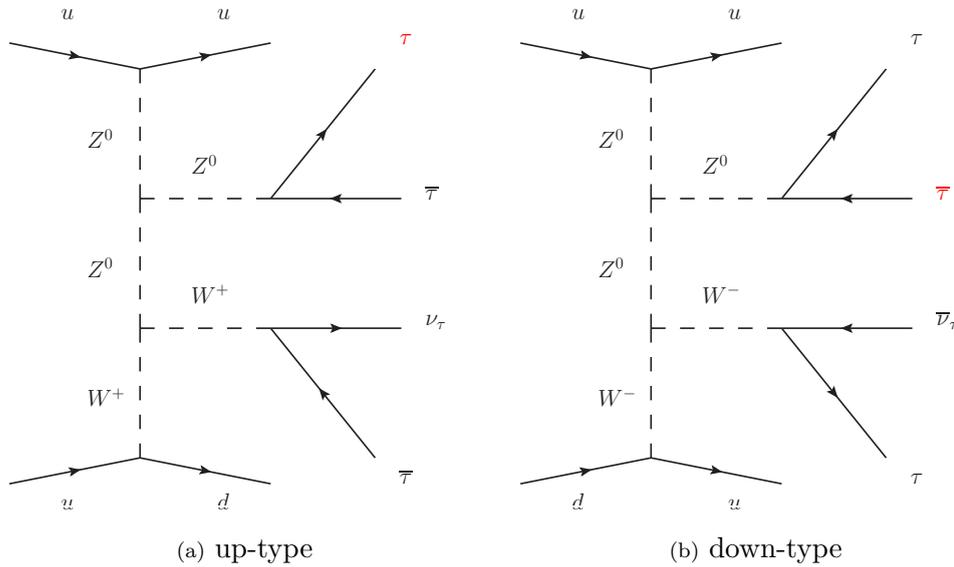
FIGURE 6.4: VBF SM production of 3 τ. Feynman diagrams show backgrounds for the vector boson fusion like sign di-τ_h supersymmetry search, where loosing one opposite sign τ (marked in red) due to failing selection thresholds or reconstruction criteria (e. g. falling out of the tracker coverage) allows for the event to enter the selection.

For this effect to occur, one of the τ leptons has to be out of the selection acceptance or tracker acceptance. Additionally, half of the already rare events with transverse momentum imbalance will be lost due to the wrong sign τ_h being reconstructed, falling into the opposite sign selection, instead. As this is, like the first group, a purely electroweak process, an accurate modeling by the simulation is expected, but the probability of such events passing selections is very small. In conclusion, this type of events is also of minor importance.

Finally, the lowest probability per individual event can be assigned to the misreconstruction of a jet as a τ_h in the order of percent chance to happen per jet (see Sec. 6.2.4

for details). On the other hand, the jet production cross section is the highest of all processes and outdoes the theoretical cross section of the VBF SUSY production by several orders of magnitude. Some example Feynman diagrams are shown in Fig. 6.5.
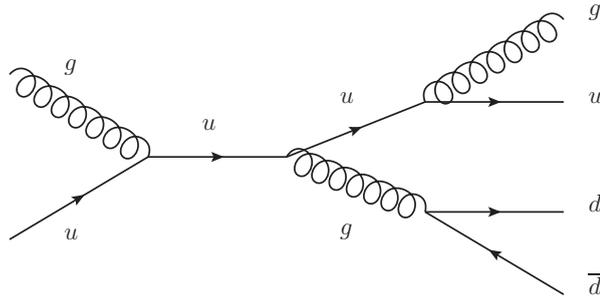
(a) An example of 4 jet QCD production with final state radiation
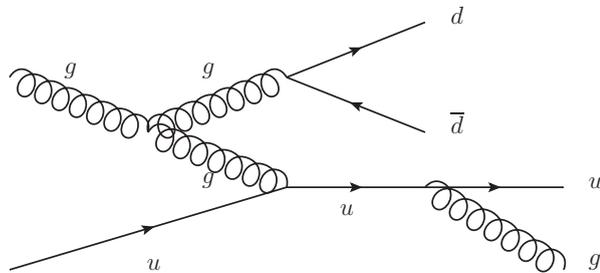
(b) An example of 4 jet QCD production with initial state radiation

FIGURE 6.5: Example Feynman diagrams of QCD backgrounds for the vector boson fusion like sign di-τ_h supersymmetry search, where two jets need to be misreconstructed as two like sign τ_h.

Even the chance for two jets from inclusive QCD production with at least four jets is, therefore, dominating the backgrounds in the like sign channel, as the misreconstructed jets (τ_h^{fake}) have random charges, at least according to simulation. This can be motivated by the realization that any jet faking a τ_h has more than one charged component and there is charge conservation. Randomly choosing one isolated charged hadron that leaves an entry in the ECAL will also randomly determine the charge of the τ_h^{fake}. Accordingly, one can expect half of all events with four jets out of which two are τ_h^{fake} to enter the like sign di-τ_h selection.

A hybrid are the W+jets, semi-leptonic top or single top cases, in which one real τ_h is present and only one jet needs to be misreconstructed as τ_h, in order to pass selections. Some examples are shown in Fig. 6.6.
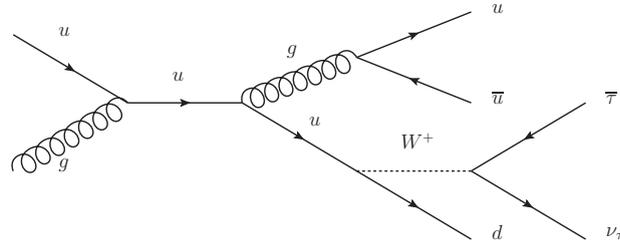
(a) SM W + 3 jets production: One jet needs to be misreconstructed as τ_h
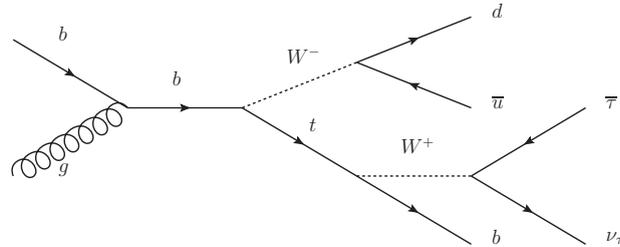
(b) SM single top production: One jet needs to be misreconstructed as τ_h, most likely the b-jet, so the b-tag does not need to be failed
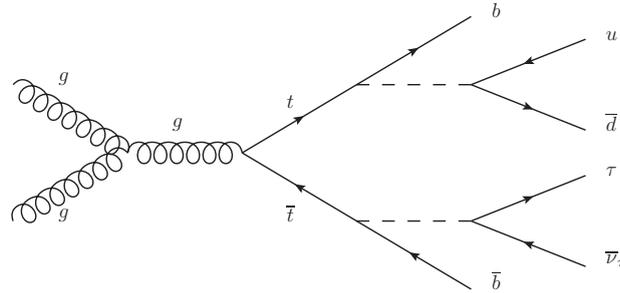
(c) Semi-leptonic SM top production: To enter selection, one of the b-jets has to fail b-tagging while the second one is misreconstructed a τ_h

FIGURE 6.6: Feynman diagrams of QCD backgrounds for the vector boson fusion like sign di-τ_h supersymmetry search, where one jet needs to be misreconstructed as a fitting sign τ_h.

Semi-leptonic top and single top are, again, suppressed by requiring no b tags in the event.

What remains as the dominant contribution for like sign events is QCD with misreconstructed jets. The subdominant contribution is likely to be W+jets with one real τ_h and one misreconstructed jet.

To estimate the QCD contamination of the signal region, a data-driven background estimation based on an ABCD method is developed. An ABCD method works by sequestering a selection of data into four subregions, out of which one is the signal region A.

In the two-dimensional plane, each axis represents one inversion argument with respect to the signal region. In this case, one axis is the requirement of the degree of τ isolation, the other axis is the requirement of two jets passing or failing all VBF selection criteria. Therefore, a region B could be tight isolation, failing VBF requirements. Meanwhile, region C and D would require a non-tight isolation and passing and failing the VBF requirements, respectively. If τ isolation and VBF requirements are not correlated, then the ratio of events passing and failing the VBF selection is independent of τ isolation and the relation $A = B \cdot \frac{C}{D}$ holds, constituting the ABCD method.

According to the previously explained ABCD method, four exclusive isolation regions are defined:

TT Two tightly isolated τ_h

TM_I One tightly isolated τ_h and one τ_h failing tight isolation and passing medium and/or loose isolation criteria (passing medium isolation necessarily requires passing loose isolation, as well)

MM_I One medium isolated τ_h failing tight isolation and passing medium and loose isolation, as well as one τ_h failing tight isolation and passing medium and/or loose isolation criteria

LL Two τ_h passing loose isolation and failing medium and tight isolation

Furthermore, failing any of the three VBF requirements specified in Sec. 6.1.1 spans up a second dimension called VBF inversion ($\overline{\text{VBF}}$).

Taking these definitions into account, the signal region (SR) and seven control regions (CR) for like sign are defined, as well as their opposite sign counterparts, as shown in Fig. 6.8.

	LS		OS	
isolation	VBF	$\overline{\text{VBF}}$	VBF	$\overline{\text{VBF}}$
TT	SR	CR2	OS SR	OS CR2
TM_I	CR3	CR4	OS CR3	OS CR4
MM_I	CR5	CR6	OS CR5	OS CR6
LL	CR7	CR8	OS CR7	OS CR8

TABLE 6.8: Control region (CR) and signal region (SR) definitions with vector boson fusion selection (VBF), their inversion $\overline{\text{VBF}}$, like sign (LS) regions, opposite sign (OS) regions, and several different isolation criteria.

The contamination of QCD expected in the signal region $N_{\mathrm{SR}}^{\mathrm{QCD}}$ is extrapolated from CR2:

$$N_{\mathrm{SR}}^{\mathrm{QCD}} = (N_{\mathrm{CR2}}^{\mathrm{data}} - N_{\mathrm{CR2}}^{\overline{\mathrm{QCD}}\ \mathrm{MC}}) \cdot \frac{\epsilon^{\mathrm{VBF}}}{1 - \epsilon^{\mathrm{VBF}}} \tag{6.1}$$

$N_{\mathrm{SR}}^{\mathrm{QCD}}$ is the QCD background expectation in the signal region, $N_{\mathrm{CR2}}^{\mathrm{data}}$ is the number of data events in control region 2 (see Tab. 6.8), $N^{\overline{\mathrm{QCD}}\ \mathrm{MC}}$ the number of expected non-QCD backgrounds in control region 2 (taken from simulation) and ϵ^{VBF} is the VBF-efficiency, defined in Eq. 6.2.

$$\epsilon^{\mathrm{VBF}} = \frac{N_{\mathrm{CR_{VBF}}}^{\mathrm{data}} - N_{\mathrm{CR_{VBF}}}^{\overline{\mathrm{QCD}}\ \mathrm{MC}}}{N_{\mathrm{CR_{VBF}}}^{\mathrm{data}} - N_{\mathrm{CR_{VBF}}}^{\overline{\mathrm{QCD}}\ \mathrm{MC}} + N_{\mathrm{CR_{\overline{VBF}}}}^{\mathrm{data}} - N_{\mathrm{CR_{\overline{VBF}}}}^{\overline{\mathrm{QCD}}\ \mathrm{MC}}} \tag{6.2}$$

Here, $\mathrm{CR_{VBF}}$ are all odd numbered control regions with VBF requirements applied and $\mathrm{CR_{\overline{VBF}}}$ their respective counterparts failing any of the VBF selection criteria ($\Delta\eta(\mathrm{jet}, \mathrm{jet})$, $\mathrm{M}^{\mathrm{jet, jet}}$ or $\mathrm{sign}(\eta^{\mathrm{jet}\ 1} \cdot \eta^{\mathrm{jet}\ 2})$ as stated in Sec. 6.1.1, but still require at least two additional jets to be part of the event. The resulting control regions used are listed in Tab. 6.9 and later in Tab. 6.15 the scale factors determined from data will be shown.

VBF	$\overline{\mathrm{VBF}}$
CR3	CR4
CR5	CR6
CR7	CR8
OS CR3	OS CR4
OS CR5	OS CR6
OS CR7	OS CR8

TABLE 6.9: List of control region pairs used to calculate ϵ^{VBF}. Opposite sign control regions are also used, because the VBF selection efficiency is not expected to depend on the sign of the τ_h leptons as none of the VBF selection requirements use τ_h leptons.

The assumption of QCD domination is of paramount importance, as the degree of modeling of non-QCD backgrounds is unknown and the subtraction of simulation based background estimates is varied by $\pm 50\%$ for all control regions to estimate the "MC systematic". Even this conservative assumption of mismodeling breaks down when another background becomes similarly relevant as QCD. Such an effect would need to be covered by additional systematic uncertainty estimates and therefore needs to be checked. Another weakness of this approach is that even with six different estimates for ϵ^{VBF}

that are combined by using a weighted arithmetic mean, it cannot be checked whether the ϵ^{VBF} depends on τ_h-isolation. For each pair of control regions used for calculating the efficiency, biases by the trigger or selection due to τ_h isolation are assumed to cancel in the ratio. The spread of values over the control regions is taken into account by the statistical uncertainty of ϵ^{VBF}, but still there could be a systematic effect of requiring only tight isolation for τ_h as done for the signal region that cannot be verified or falsified in data itself, as there are not sufficient events for a signal-like sideband.

Furthermore, to increase the statistical accuracy of ϵ^{VBF}, the central selection requirement of $\not{E}_{\text{T}} \geq 30$ GeV is dropped for all control regions listed in Tab. 6.9. This could introduce a bias into the measurement of ϵ^{VBF}, as there might be a different correlation of jets passing and failing VBF selections with this requirement, e.g., due to the rapidity gap. Even if the jet energy corrections work at the forward parts of the detector, the higher jet energy resolution values there allow for more misbalanced events than in the barrel, producing artificial $\not{E}_{\text{T}}$. A requirement on the missing transverse momentum will therefore bias the selection of QCD events towards more mismeasured events that are more likely to have high rapidity gaps and therefore pass the VBF selections. The extent of such an effect can also not be checked on data, as there are not enough events to make a variation of the requirement on $\not{E}_{\text{T}}$.

These systematical problems can be solved with the simulation based approach discussed in Sec. 6.2.

The expected number of background events in the signal region is the direct number of non-QCD backgrounds from simulation, varied by $\pm 50\%$ for a MC systematic uncertainty, plus $N_{\text{SR}}^{\text{QCD}}$ from Eq. 6.1. The results are for ϵ^{VBF}:

$$\epsilon^{\text{VBF}} = 6.7 \pm 0.46 \text{ (stat) } {}^{-0.038}_{+0.022} \text{ (MC) } \% \tag{6.3}$$

and for $N_{\text{SR}}^{\text{QCD}}$:

$$N_{\text{SR}}^{\text{QCD}} = 7.15 \pm 0.92 \text{ (stat) } {}^{-0.42}_{+0.35} \text{ (MC)} \tag{6.4}$$

with $N_{\text{SR}}^{\overline{\text{QCD}}}$:

$$N_{\text{SR}}^{\overline{\text{QCD}}} = 0.83 \pm 0.079 \text{ (stat) } \pm 0.41 \text{ (MC)} \tag{6.5}$$

6.2 Validation of the QCD background in simulation

While a data-driven approach with an ABCD method in data is utilizable in order to get a prediction for the number of background events expected in the signal region, there are two advantages of a simulation-based approach. On the one hand, one can ascertain that the control regions contain only QCD. On the other hand, one can provide a closure test for the data-driven prediction, especially regarding the universality of the VBF efficiency with respect to a loosening of the τ_h isolation or a relaxation of the $\not{E}_T$ requirement, as stated in Sec. 6.1.3.

In order to achieve the goal to ensure QCD purity in the control regions, a straightforward use of the simulation is not possible, due to the low probability of jets to fake τ_h with any identification (ID), be it tight (T), medium (M), or loose (L). These low rates decrease the number of simulated events with two reconstructed τ_h leptons passing all selection requirements below any reasonable amount for direct usage.

What can be done, is to estimate the probability of one jet to be misidentified as a τ_h lepton ($\mathcal{O} \approx \frac{1}{100}$) of a given exclusive isolation (see Tab. 6.8 in Sec. 6.1.3 for definitions), instead. Using all simulated events with four or more jets, one can determine the chance of each such event to be reconstructed as one event with two τ_h leptons and at least two jets, translating two jet objects to become τ_h^{fake} objects (see Sec. 6.2.1). This way, a meaningful study of systematic uncertainties of the background estimation method on data becomes possible and the prerequisite assumptions of QCD purity and stability of the VBF efficiency with respect to τ_h^{fake} isolation can be tested.

Finally, the results of the application of this method and thereby estimated systematic uncertainties will be shown (see Sec. 6.3). As well, possible future data-driven improvements for this method will be outlined (see Sec. 6.3.4).

6.2.1 Redefining jets as τ_h

In a simulated event with no real τ_h leptons on generator level, the chance of reconstructing two jets as τ_h^{fake} leptons is needed to be calculated, in order to reuse all events with four or more jets for gaining statistical accuracy. Given two jets having the same probability to fake, e.g., a τ_h lepton passing tight isolation criteria (T), an event weight $w_{\text{event}}^{\text{TT}}$ is calculated as the probability of the event to fake exactly two τ_h leptons:

$$w_{\text{event}}^{\text{TT}} = \sum_{i=0}^{N} P(\text{T}|\text{jet}_i) \left(\sum_{j=0;\ i \neq j}^{N} P(\text{T}|\text{jet}_j) \left[\prod_{k=0;\ k \neq i,j}^{N} (1 - P(\text{T}|\text{jet}_k)) \right] \right) \quad (6.6)$$

N is the number of eligible jets. The conditional probability $P(\text{T}|\text{jet}_i)$ given a jet_i is more precisely defined in Eq. 6.7.

$$P(\text{T}|\text{jet}_i) = P(\text{tight ID}|\text{jet}_i) \cdot P(p_T^{\tau_h^{fake}} > 45 \text{ GeV}|\text{jet}_i) \cdot \epsilon^{\text{trigger}}(\text{ID}) \qquad (6.7)$$

$P(\text{tight ID}|\text{jet}_i)$ is the probability of the jet with index i to pass a tight τ_h-lepton isolation according to the efficiency out of a map described in Sec. 6.2.4. $P(p_T^{\tau_h^{fake}} > 45 \text{ GeV}|\text{jet}_i)$ is the acceptance correction regarding the selection threshold of 45 GeV for τ_h leptons and is described in Sec. 6.2.2. Finally, the chance of a trigger leg firing due to a τ_h^{fake} object of a given isolation requirement, is determined according to Sec. 6.2.5. This effect is taken into account by the trigger acceptance correction $\epsilon^{\text{trigger}}(\text{ID})$.

For control regions with two different isolation requirements for the τ_h^{fake} objects, e.g., CR3 or CR4 (see Tab. 6.8) requiring one reconstructed τ_h lepton to fulfill tight isolation criteria (T), while the second τ_h^{fake} lepton may have medium or loose isolation (M$_\text{I}$ for medium inclusive), the event weight $w_{\text{event}}^{\text{TM}_\text{I}}$ in Eq. 6.8 is assigned.

$$w_{\text{event}}^{\text{TM}_\text{I}} = \sum_{i=0}^{N} P(\text{T}|\text{jet}_i) \left(\sum_{j=0; \ i\neq j}^{N} P(\text{M}_\text{I}|\text{jet}_j) \left[\prod_{k=0; \ k\neq i,j}^{N} (1 - P(\text{T}|\text{jet}_k) - P(\text{M}_\text{I}|\text{jet}_k)) \right] \right) \tag{6.8}$$

Since the identification criteria are defined exclusively, Eq. 6.9 holds true:

$$P(\text{M}_\text{I}|\text{jet}_i) = P(\text{M}|\text{jet}_i) + P(\text{L}|\text{jet}_i) \qquad (6.9)$$

The object transverse momentum translation function for an inclusive subsample of all isolations needs to take into account the relative mixture of isolation criteria involved, because different isolation requirements can lead to different response functions of the corresponding τ_h^{fake} objects $\mathcal{R} = \frac{p_T^{\tau_h^{\text{fake}}}}{p_T^{\text{jet}}}$. The relative probability of a jet being reconstructed as a τ_h^{fake} with any of the different isolation criteria under consideration, including the respective selection and trigger acceptance corrections, is then used as a weight. A weighted average of the different contributions for the $p_T^{\text{jet}} \to p_T^{\tau_h^{\text{fake}}}$ translation factor $x(\text{ID}|\text{jet}_i)$ (see Sec. 6.2.3) is used to get the correct average translation factor for

the particular probability mixture of isolations for that jet, as described in Eq. 6.10.

$$x(\mathrm{M_I}|\mathrm{jet}_i) = \frac{P(\mathrm{M}|\mathrm{jet}_i) \cdot x(\mathrm{M}|\mathrm{jet}_i) + P(\mathrm{L}|\mathrm{jet}_i) \cdot x(\mathrm{L}|\mathrm{jet}_i)}{P(\mathrm{M}|\mathrm{jet}_i) + P(\mathrm{L}|\mathrm{jet}_i)} \tag{6.10}$$

After calculating the event weight, the jets faking τ_h leptons are randomly chosen according to their individual fake probability out of all jets with $P > 0$ in the order of the tighter isolation to be chosen first, if both isolations differ. The same events are reused for all τ_h^{fake}-isolation regions, as seen in Tab. 6.10. The last necessary step is to transform $p_T^{\mathrm{jet}} \to p_T^{\tau_h^{\mathrm{fake}}}$ using scale factors explained in Sec. 6.2.3 and assign the individual τ_h^{fake} object a reasonable mass, as explained in Sec. 6.2.6.

	$T_1^{fake/real}$	$M_1^{fake/real}$	$L_1^{fake/real}$
T_2^{fake}	SR/CR2	CR3/CR4	CR3/CR4
M_2^{fake}	CR3/CR4	CR5/CR6	CR5/CR6
L_2^{fake}	CR3/CR4	CR5/CR6	CR7/CR8

TABLE 6.10: Table showing assignment of all possible τ_h-isolation combinations into control regions (CR) and signal region (SR) as in Fig. 6.8. Despite the index, p_T-ordering is decided after dicing and assignment to a control region. The superscript *fake/real* describes fakes for events containing no generator τ_h or instead for events with one generator level τ it describes one real τ_h.

For samples with one real τ_h, like W+jet and single top, dicing is also applied to get more accurate estimates. Here, only one jet is diced to become a τ_h^{fake} object and the event weight is calculated according to Eq. 6.11.

$$w_{\mathrm{event}}^{\mathrm{T^{fake}T^{real}}} = \epsilon^{\mathrm{trigger}}(\mathrm{T^{real}}) \cdot \sum_{i=0}^{N} P(\mathrm{T}|\mathrm{jet}_i) \left[\prod_{j=0;\ j \neq i}^{N} (1 - P(\mathrm{T}|\mathrm{jet}_j)) \right] \tag{6.11}$$

The assignment, whether zero, one, or two τ_h^{fake} objects need to be diced (N^{diced}), is done event by event, taking into account the number of generator level τ (N^{gen}) in that respective event, as described in Eq. 6.12.

$$N^{\mathrm{diced}} = 2 - N^{\mathrm{gen}} \tag{6.12}$$

It is not explicitly required that these real generator τ leptons pass the selection requirements or decay hadronically.

All parts of Eq. 6.7 are derived with respect to jet parton flavor:

- Light quarks (up, down or strange)

- Charm quarks

- Bottom quarks

- Gluons

- Unmatched contributions (e.g., pileup jets)

This is done, in order to avoid any biases by the choice of the sample(s) the recipe is derived on and/or applied to.

6.2.2 Selection acceptance correction

In this section, the selection acceptance correction $P(p_T^{\tau_h^{\mathrm{fake}}} > 45\text{ GeV}|\mathrm{jet}_i)$ out of Eq. 6.7 is explained in detail.

The transverse momentum of a jet does not directly translate into the transverse momentum of a τ_h it may be misreconstructed as. Therefore, requiring $p_T^{\tau_h} \geq 45$ GeV does not mean an equal requirement in terms of p_T^{jet}. Such a statement equals a distribution of a response $\mathcal{R}$ with a nonzero width, introduced in Eq. 6.13.

$$\mathcal{R} = \frac{p_T^{\tau_h^{\mathrm{fake}}}}{p_T^{\mathrm{jet}}} \tag{6.13}$$

The fraction of such a distribution where a given jet transverse momentum translates into a τ_h^{fake} transverse momentum exceeding selection requirements, equals the selection acceptance correction $P(p_T^{\tau_h^{\mathrm{fake}}} > 45\text{ GeV}|\mathrm{jet}_i)$.

In principle, one could account for this effect by making an explicit selection requirement of $p_T^{\tau_h^{\mathrm{fake}}} > 45$ GeV when deriving efficiency maps out of Sec. 6.2.4. Here, in order to be able to quantify the effect, a factorized approach was chosen. By counting all events in a given range of p_T^{jet} that pass or fail the selection requirement of $p_T^{\tau_h} \geq 45$ GeV on reconstructed τ_h matched within $\Delta R < 0.1$ within simulated events and then dividing

this number by all events with matched τ_h^{fake}, the acceptance displayed in Fig. 6.7 is determined.

(a) Inclusive flavor

(b) Tight isolation by flavor

(c) Medium isolation by flavor

(d) Loose isolation by flavor
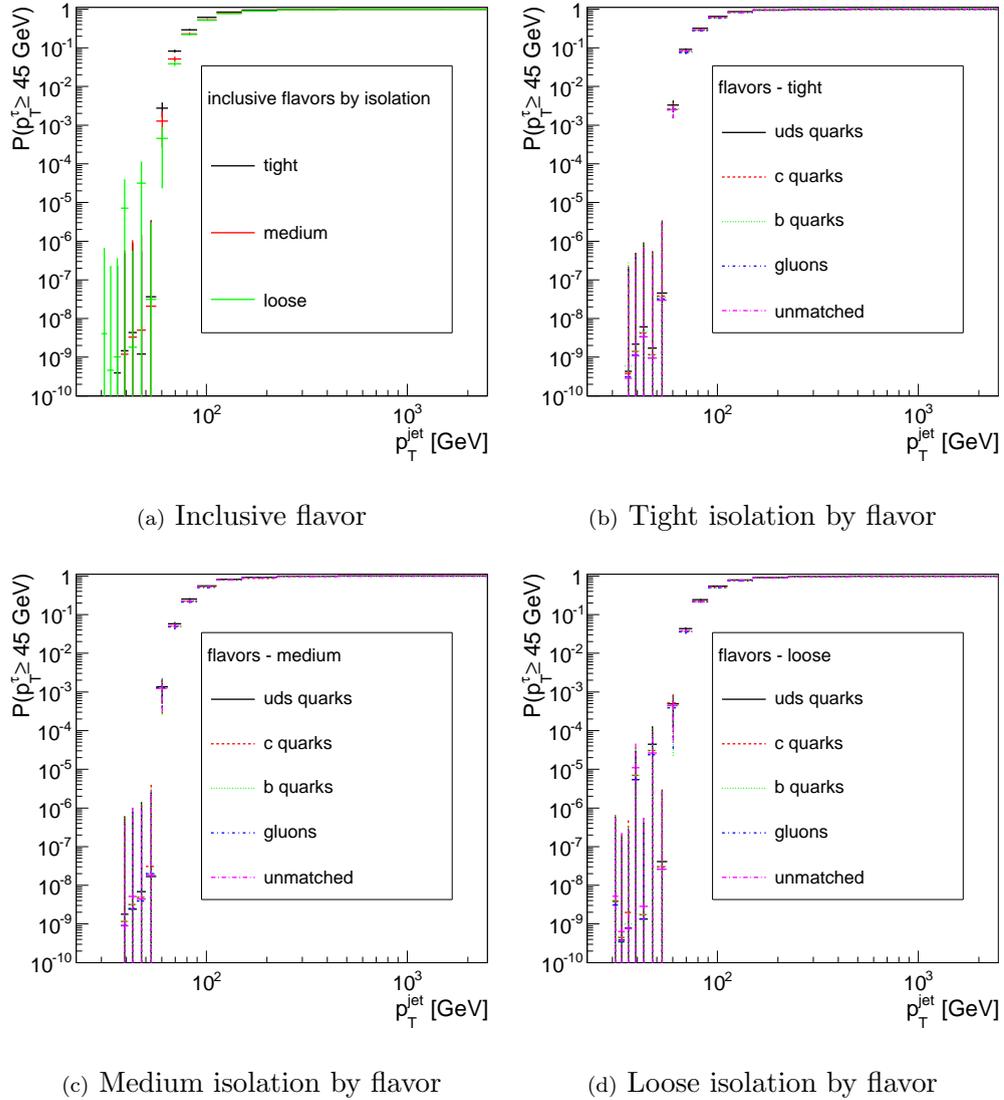
FIGURE 6.7: Selection acceptance correction for diverse flavors and isolations. The general shape is very similar for all isolations, but there are differences outside of the statistical uncertainties, especially at low momenta. In general, low jet momenta are very unlikely to pass the τ_h transverse momentum requirement. Regarding flavor, there is no observation of any differences for any of the isolations.

In the end, the selection acceptance correction asymptotically approaches full efficiency at very high jet momenta, showing slight differences for each isolation in the 100 GeV jet transverse momentum region. Due to the exponentially falling jet transverse momentum spectrum, even jets suppressed by this correction have substantial contributions. Differences between isolations, like for $p_T^{\text{jet}} \approx 90$ GeV jets or at very high momenta, do

translate into shape changes per isolation. Such changes are shown for the example of the second to leading τ_h^{fake} transverse momentum in Fig. 6.8.
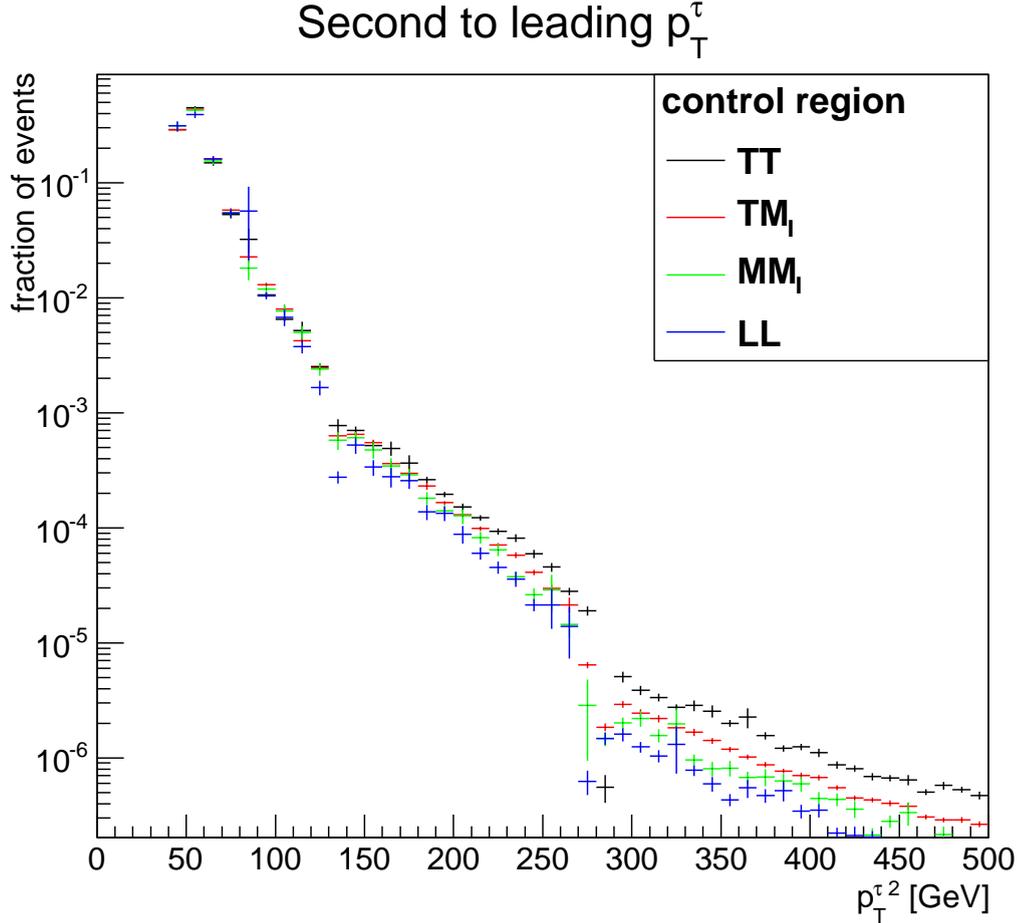
FIGURE 6.8: Shape changes for different control region compositions (T is tight, M is medium, L is loose, subscript I means that lower isolations are allowed) are shown for the second to leading τ_h^{fake}. At high transverse momenta, the largest differences are observed.

6.2.3   Jet to τ_h transverse momentum translation

A jet and the fake τ_h it might end up being reconstructed as, share some, but not all of their particle flow components. Moreover, the jet is subject to energy corrections while the τ_h^{fake} is not. The direct consequences are slight deviations in the direction of the two objects and major differences in the transverse momentum reconstructed. Using the response formulation from Eq. 6.13, the momentum of the originating jet is multiplied with the average of the response distribution, in order to obtain a value for the transverse momentum of the τ_h^{fake}.

The direct mean of all entries per bin of p_T^{jet} passing all selection requirements is used to avoid shifts of the arithmetic mean by different binnings of the response distribution. These average translation factors (transmutation factors) are applied on the p_T^{jet} to calculate the corresponding $p_T^{\tau_h^{\text{fake}}}$ and are shown in Fig. 6.9.

(a) Inclusive flavor

(b) Tight isolation by flavor

(c) Medium isolation by flavor

(d) Loose isolation by flavor
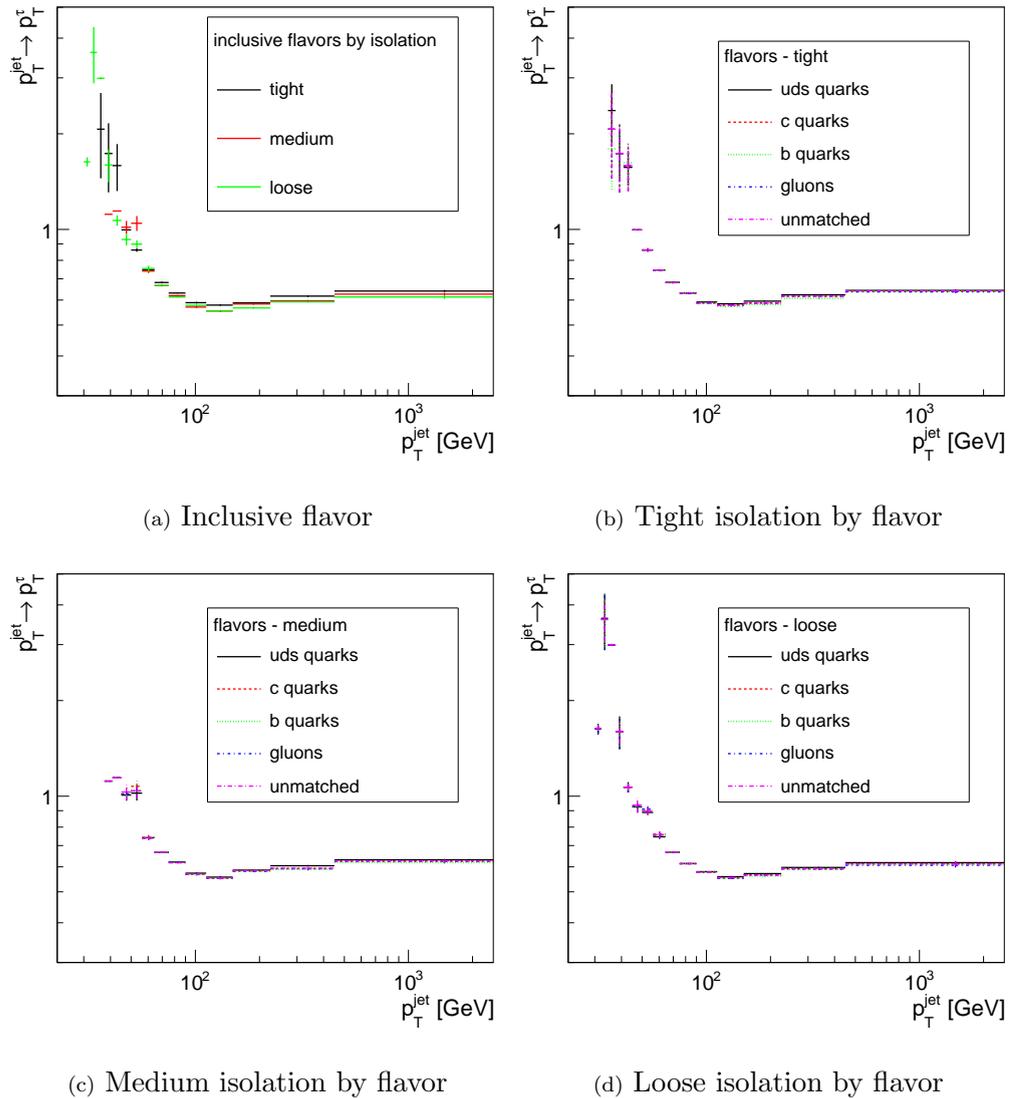
FIGURE 6.9: Displayed is the transmutation factor for different isolations. Only minor differences are observed. The flavor of the originating jet does not seem to have any impact on the reconstructed object's transverse momentum on average.

Two effects can be observed. On the one hand, at very low momenta the selection requirement of $p_T^{\tau_h} > 45$ GeV is excluding lower response values from the visible response distribution. This forces the transmutation factors to higher values while at the same time reducing the overall likelihood of such an event being selected, as shown in Sec. 6.2.2.

On the other hand, the jet energy correction factors are lowest at very high momenta. Thus, the difference of the corresponding τ_h^{fake} objects at such high momenta are most similar to the jets in that region of phasespace. That effects a rise of the transmutation factor $p_T^{\text{jet}} \to p_T^{\tau}$ in that region.

One would expect low reconstructed particle multiplicities with a large spread at low momentum necessitating that low fractions of jet energy are reconstructed as a τ_h^{fake}. This would motivate a turn-on curve of the transmutation factor reaching a plateau at high jet momenta, when the statistical spatial distribution of jet particles would begin to dominate with respect to the momentum distribution among particles. Therefore, a counterintuitive effect of this transmutation factor distribution is that the lower edge of the transverse momentum distributions of the τ_h^{fake} at $45 \text{ GeV} \leq p_T^{\tau_h^{\text{fake}}} < 60 \text{ GeV}$ consists of the $30 \text{ GeV} \leq p_T^{\text{jet}} < 100 \text{ GeV}$ region in terms of jet transverse momentum, as shown in the next section.

6.2.4 Choice of parametrization for fake probabilities

Choosing a parametrization to emulate the properties of the τ_h identification (ID) is a matter of finding a meaningful description that is different enough from the τ_h ID. The τ_h ID yields not a single event after requiring the central and the VBF selections, no matter the isolation (see Tab. 6.14). In the end, this section is the effort to maximize the statistical precision while minimizing the systematic deviation from the predictions of the actual τ_h ID, because the VBF topologies examined are very small portions of phasespace making it imperative to find as general a parametrization as viable.

As jets are objects clustered out of many parts, charged and not charged, out of the hadronization of color charged particles, the natural choice of parameters corresponds to the amount of particles, the charge distribution of particles and the particle density in the jet. Furthermore, one can check whether detector effects play a role. Following this logic, four variables were considered:

N^q Number of charged particles: Corresponds to the amount of charged particles. It is approximately proportional to the amount of particles in the jet and therefore to the charge distribution.

F^q Fraction of charged particle momentum: The momentum of all charged jet components, relative to the overall jet momentum. It is proportional to the momentum distribution of charged and neutral particles and, to some extent, to the amount of charged particles. Extreme values typically correlate with low amounts of charged particles.

p_T^{jet} Jet transverse momentum: The higher the transverse momentum, the larger the hadronization phasespace, leading to larger amounts of particles involved. Furthermore, the boost of the initial parton translates to smaller opening angles of an imaginative cone surrounding all jet constituents, giving indirect information about the particle density.

$|\eta|$ Jet pseudorapidity: If there are correlations between detector effects and τ_h ID, then this variable will show a dependency of the misreconstruction probability. Otherwise, it is expected to be flat with respect to this variable.

In this subsection only, a study with less complexity than the full background reproduction of two τ_h^{fake} objects is performed, to find the parametrization. The probability of a jet to be misreconstructed as a τ_h lepton is here defined as the ratio of a numerator, consisting of all jets within $|\eta| \leq 2.2$ and with $p_T^{\text{jet}} \geq 30$ GeV matched within $\Delta R < 0.1$ to reconstructed τ_h within $|\eta| \leq 2.1$ passing tight isolation criteria, and a denumerator, consisting of all jets within $|\eta| \leq 2.2|$ and with $p_T^{\text{jet}} \geq 30$ GeV. All QCD samples are taken into account for this efficiency map and weighted according to their cross section and luminosity before adding the respective numerator and denumerator two-dimensional distributions with all six unique permutations of the former listed variables. Then, the ratio is taken. The resulting efficiency maps are shown in Fig. 6.10.

(a) N^q vs F^q

(b) N^q vs p_T^{jet}

(c) N^q vs $|\eta|$

(d) F^q vs p_T^{jet}

(e) F^q vs $|\eta|$

(f) p_T^{jet} vs $|\eta|$
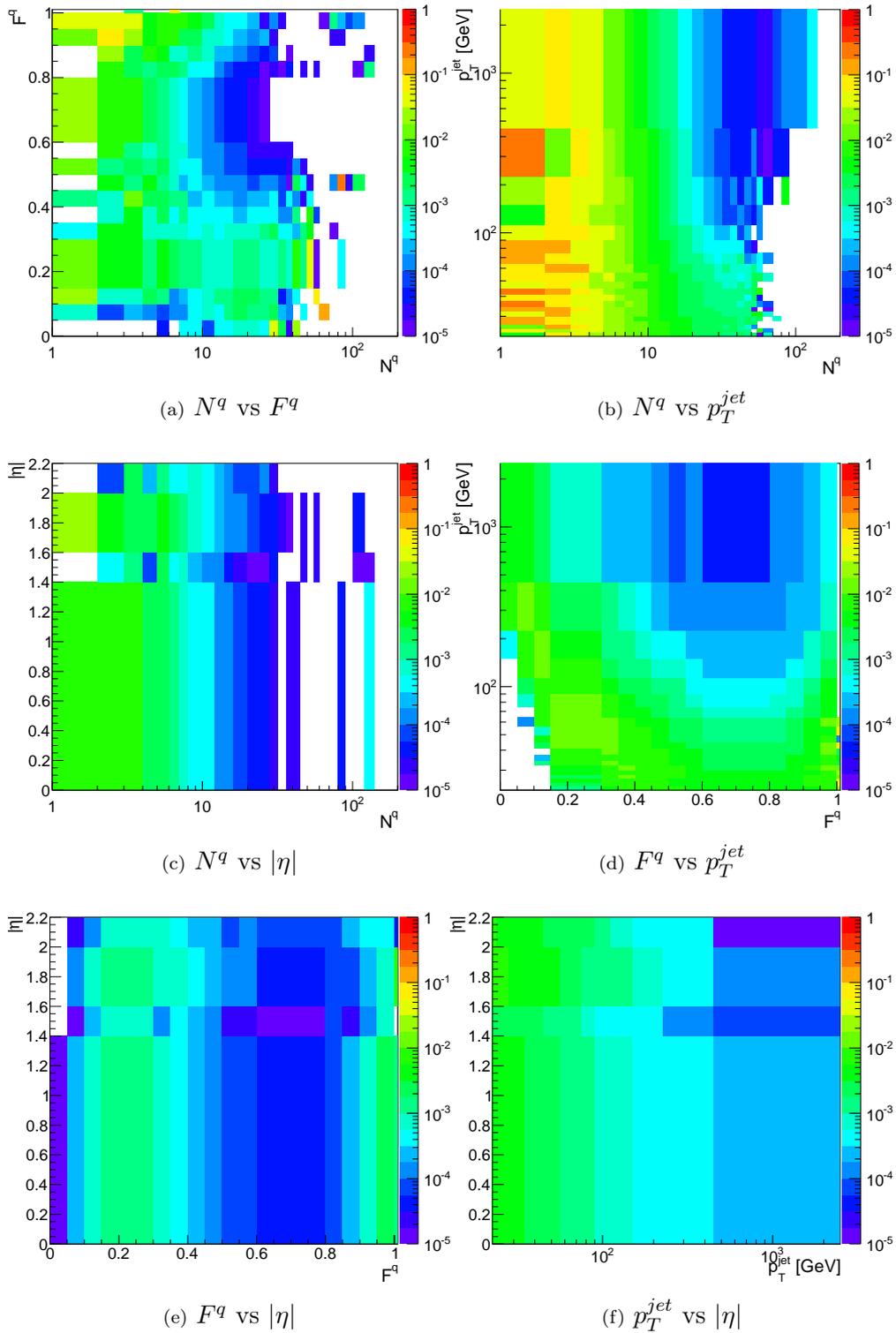
FIGURE 6.10: Two-dimensional efficiency maps for jets reconstructed as τ_h leptons passing tight isolation criteria in different parametrizations. Determined on QCD samples, only.

Instead of using the full dicing procedure explained in Sec. 6.2.1, for the sake of a closure test, the full selections as described in Sec. 6.1.3 is performed, but for the full requirement of two τ_h leptons and the trigger requirement. Instead, one jet is randomly chosen and redefined as a τ_h changing only its transverse momentum as described in Sec. 6.2.3. Then, two different variants are produced:

1. One τ_h^{fake} object inherent to the event (taken from simulation, directly) in addition to the randomly chosen jet perform as the two τ_h leptons required by the selection criteria. This variant is called the baseline as it consists of objects taken directly from the simulation, only.

2. There is no τ_h-reconstruction performed. Instead, one jet is diced to become a τ_h lepton according to the efficiency map under scrutiny. A second jet is randomly chosen to perform as a τ_h lepton. In essence, one of the objects is taken from dicing depending on the chosen parametrization and can therefore be compared to the baseline to judge its performance.

Comparing these two variants in various observables allows to have a study on which variables suffice to reproduce the input shape best or at all. In the future, a thorough use of this method may enable one to determine systematic uncertainties on the predicted shapes of diverse observables. This has not been done, in this case, as the analysis this background validation is performed for consists of cut and count, only, not using any shape information beyond the verification of the dominant presence of QCD background events.

Using this definition, the level of closure in Fig. 6.11 for the transverse momentum of the leading and second to leading jets posing as τ_h leptons is reached.
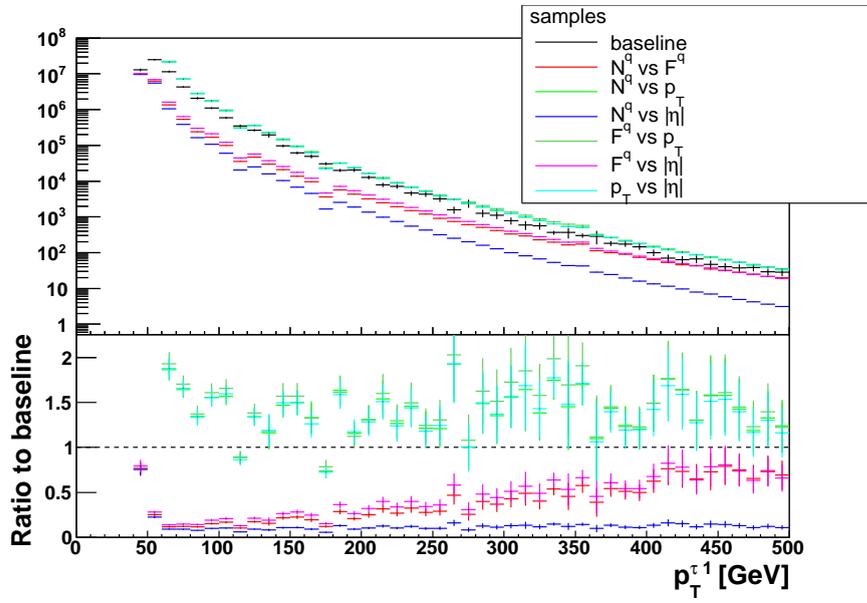
(a) $p_T^{\tau\,1}$
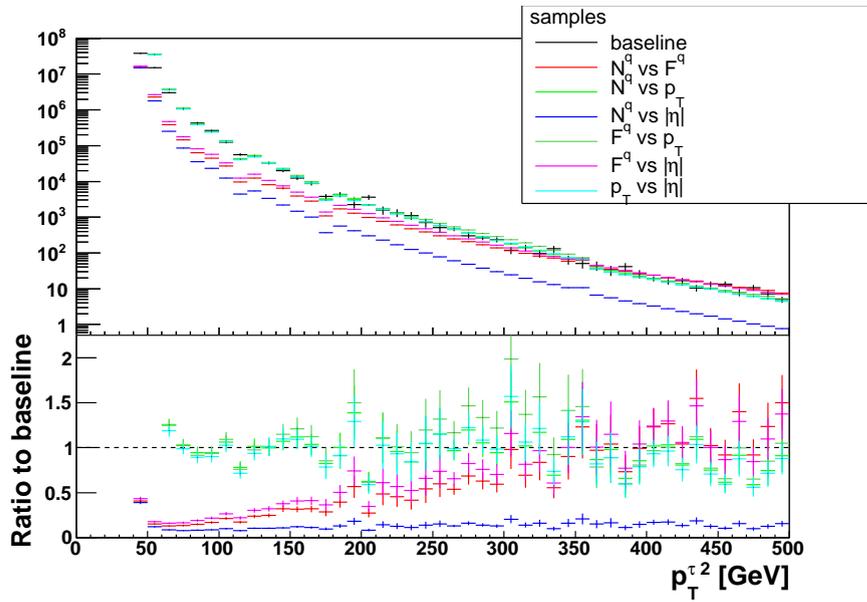
(b) $p_T^{\tau\,2}$

FIGURE 6.11: Closure in p_T^τ of diverse parametrizations for the probability of a jet to be misreconstructed as a τ_h lepton, not taking into account biases due to τ_h-selection acceptance effects.

Although the second to leading τ^{fake} shows good closure for some samples in the range of 70 to 500 GeV, the closure is very bad at low transverse momenta where the prediction is overshooting the reference by several orders of magnitude. The effect of the transverse momentum requirement of selected τ_h is not taken into account, so far. It is expected that low momentum jets have a low chance of being reconstructed as a high momentum

τ_h lepton. Taking this into account by either using a selection acceptance correction as described in Sec. 6.2.2 or by deriving the efficiency maps only with τ_h^{fake} objects passing a $p_T^{\tau_h} \geq 45$ GeV selection in the numerator yields a much better agreement, save for the lower edge of the distribution, shown in Fig. 6.12.
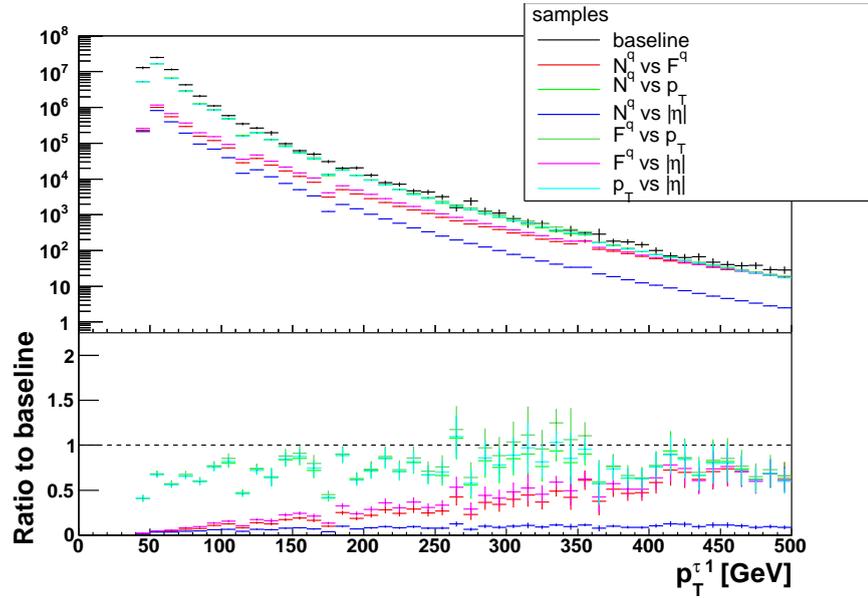
(a) $p_T^{\tau\,1}$
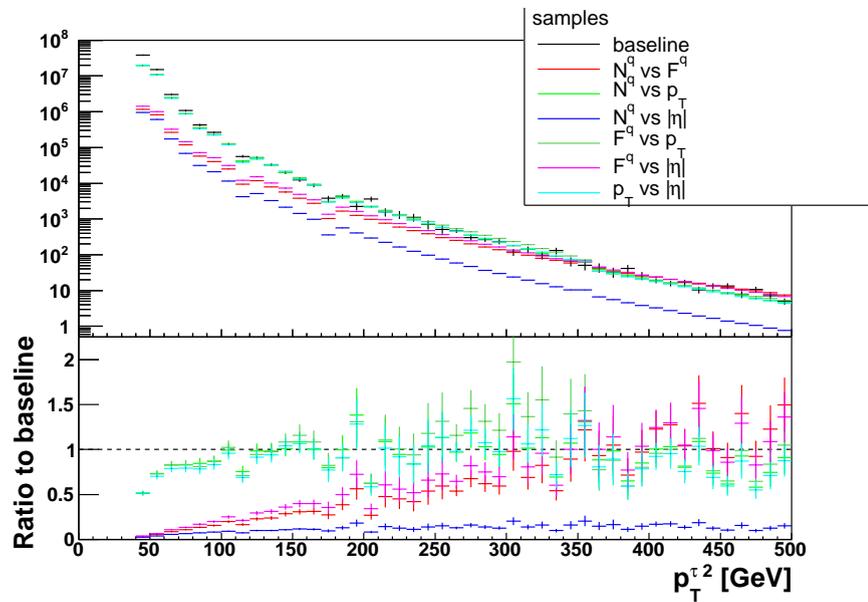
(b) $p_T^{\tau\,2}$

FIGURE 6.12: Closure in p_T^τ of diverse parametrizations for the probability of a jet to be misreconstructed as a τ_h lepton, taking into account biases due to τ_h-selection acceptance effects. Some residual discrepancy at low transverse momenta is observed.

The reason why there is an underprediction of low momentum τ_h is due to the $p_T^{\text{jet}} \to p_T^\tau$ translation described in Sec. 6.2.3. While the definition of the selection acceptance correction correctly takes into account all jets matched to τ_h^{fake} passing $p_T^\tau \geq 45$ GeV in a given bin of p_T^{jet}, this is not true for the momentum translation. Fig. 6.13 shows the $p_T^{\text{jet}} \to p_T^\tau$ translation and the selection requirement of $p_T^\tau \geq 45$ GeV.
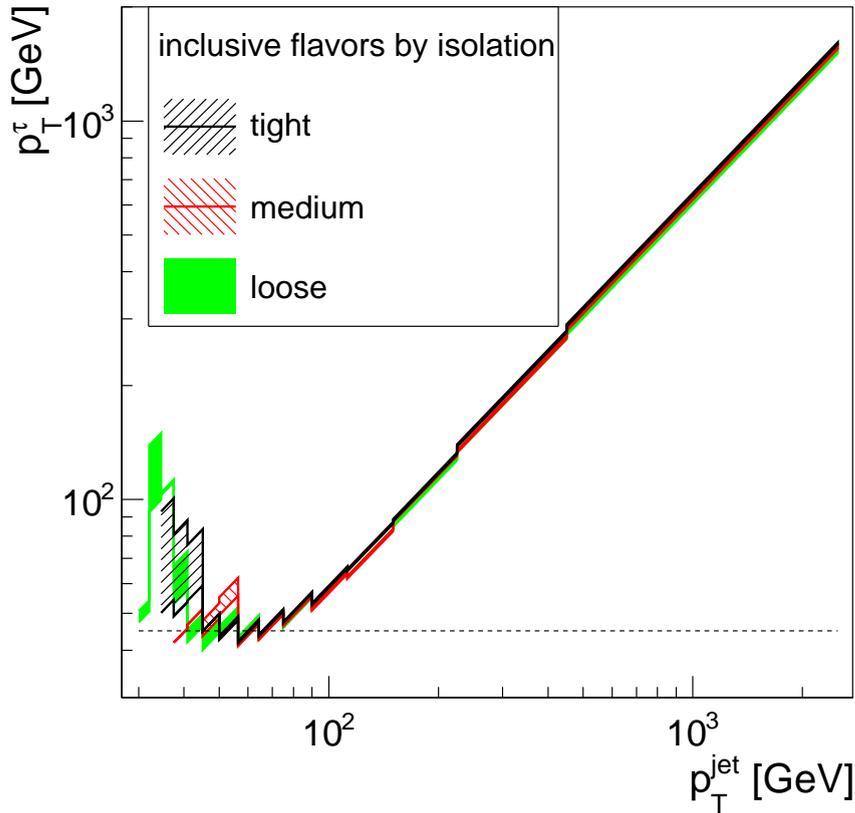
FIGURE 6.13: The jet and corresponding τ_h^{fake} momenta are shown in the same binning as the $p_T^{\text{jet}} \to p_T^\tau$ translation factors. The dotted line indicates the $p_T^\tau \geq 45$ GeV requirement.

The translation factors correspond to the accessible part of a gaussian core plus tail $p_T^\tau/p_T^{\text{jet}}$ response distribution. At high momenta, most of the response distribution is accessible. Only very low responses are forbidden due to the $p_T^\tau \geq 45$ GeV requirement. Therefore, there is a good description of the translation factor with low uncertainties. At medium jet momenta, the selection requirement starts forbidding access to the core region of the response. This necessitates that the mean of the steeply falling flank of the response distribution that is still accessible is close to the lowest allowed value of the response. Due to the finite width of the binning in p_T^{jet}, the low momentum end of each bin in this region does not pass selection requirements, anymore. Because these events have been taken into account for deriving the probability of a jet to pass $p_T^\tau \geq 45$

GeV, there is an overall underprediction. In order to solve this problem, one could either make the bin widths infinitesimally small or try to fit the response functions and determine an overall unbinned function for the momentum translation. Neither way was feasible due to the insufficient amount of simulated events available in the critical response regions. Still, this effect explains the underprediction at low p_T^{τ} for the second to leading τ_h and a corresponding flat underprediction for the leading τ_h. As this method will mostly be used for systematic studies of different τ_h isolations and scaled comparisons of data to simulation, this systematic effect should cancel in simulation to simulation comparisons and be taken into account when interpreting second to leading τ_h momentum comparisons of data to simulation.

Finally, low jet momenta necessitate looking at the tail region, only. The response in this region is almost flat, immediately increasing the average translation factor by large values, but also decreasing the probability to pass selection requirements by several orders of magnitude as shown in Sec. 6.2.2. Therefore, these anomalous events enter the intermediate transverse τ_h momentum region in negligible amounts.

To conclude, p_T^{jet} is the variable that is absolutely necessary out of the set of considered variables, in order to describe the behaviour of τ_h^{fake} objects. This effect either means that the boost of particles and therefore the particle density in the jet is an important parameter, or that the relative mixture of quark jets to gluon jets that is correlated to the particle density is important, or a mixture of both. The relative abundance of gluon-initiated jets for QCD is strongly dependent on the jet energy, because the parton density functions prefer quarks to carry large fractions of the proton momentum, as seen in Fig. 6.14.
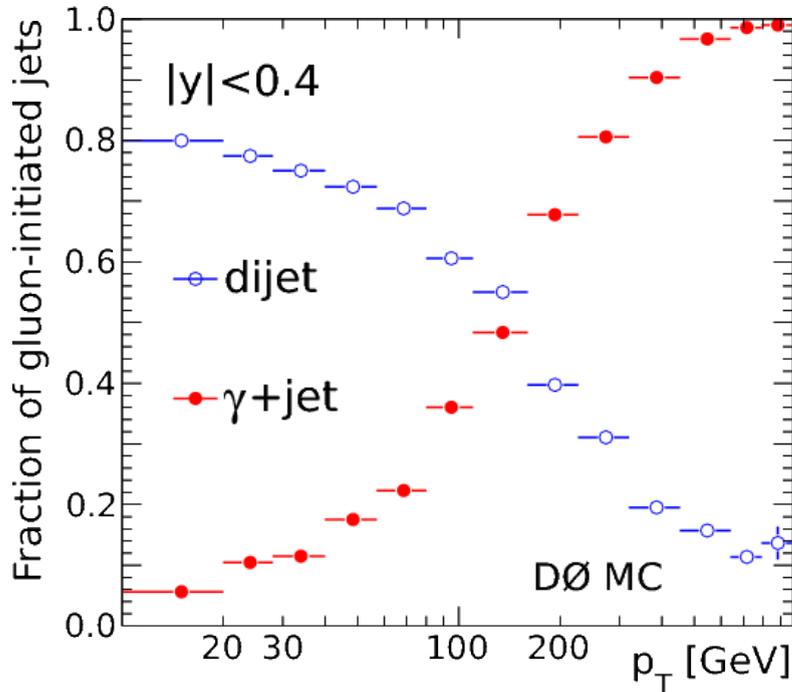
FIGURE 6.14: Taken from Ref. [78], the fraction of gluon-initiated jets with respect to jet transverse momentum in dijet events at the Tevatron is shown. There is a strong dependence of the flavor mixture in a QCD sample on the jet transverse momentum.

In addition to p_T^{jet}, the charged energy fraction F^q is taken into account, as this efficiency map excludes some jets that seem to be incapable to be misreconstructed as τ_h, unlike, e.g., the variant taking into account pseudorapidity. In jets with low momentum and a low charged energy fraction, it is not expected that any single charged particle has sufficient momentum to be reconstructed as a τ_h or that it could be isolated. Moreover, jets with low charged particle content have to rely on calorimeter based measurements. At low momenta, these objects can be expected to have a bad jet energy resolution that could by chance translate to outlying events in terms of τ_h^{fake} momentum. Therefore, excluding these jets by a proper parametrization is well motivated.

In order to calculate the probability of a jet to be misreconstructed as a τ_h^{fake} object even more thoroughly and sample independently, the numerator is redefined as any jet with $p_T^{\text{jet}} \geq 30$ GeV within $|\eta| \leq 2.6$ matched to a reconstructed τ_h lepton within $\Delta R < 0.1$ and not matched to a generator lepton within $\Delta R < 0.3$. Furthermore, the flavor of the parton producing the jet is taken into account by matching the jet to a algorithmically defined generator particle within $\Delta R < 0.3$. The denumerator, on the other hand, is redefined as all jets with $p_T^{\text{jet}} \geq 30$ within $|\eta| \leq 2.6$, again sorted by matches to respective generator particles. The classes of generator particles that are differentiated are listed

at the end of Sec. 6.2.1, but summarized in short: Light quarks, charm quarks, bottom quarks, gluons or unmatched contributions.

This results in a total of fifteen efficiency maps, five flavors times three different isolation criteria. All maps are shown in the appendix in Sec. B.1.

Regarding the differences by jet flavor, gluon jets and light quark jets do show substantial differences, as seen in Fig. 6.15.
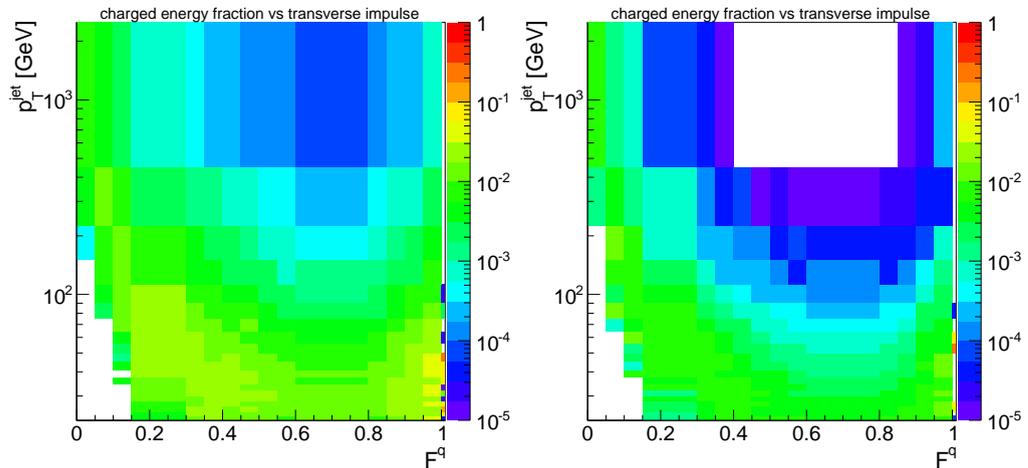
(a) Light quark jets (up, down or strange flavor)

(b) Gluon jets

FIGURE 6.15: Two-dimensional efficiency maps for jets reconstructed as τ_h leptons passing tight isolation criteria for two different jet flavors.

This behaviour can be explained by the larger color charge of gluons producing, on average, more particles per jet, as shown in Ref. [79]. The additional particles lower the chance of that jet to be misreconstructed as τ_h lepton due to a higher chance to fail isolation requirements, as can be motivated by looking at, e.g., Fig. 6.10(b), showing the fake probability in terms of the number of charged constituents of the jet. Higher numbers of constituents lower any chance of the jet to become a $p_T^{\tau_h^{\text{fake}}}$ object.

6.2.5 Trigger acceptance correction

In this section, the trigger acceptance correction $\epsilon^{\text{trigger}}(\text{ID})$ (see Eq. 6.7) will be derived. It is assumed throughout this section that reconstructed τ_h objects, whether out of jets or real τ leptons, are the dominant contribution to firing a trigger leg. The chance of any other object to do so is neglected. Furthermore, due to an insufficient number of simulated events, it is assumed that all trigger legs are fully efficient with respect to the

offline selection requirement of $p_T^{\tau_h^{\text{fake}}} \geq 45$ GeV.

On a DYToTauTau-sample, the trigger acceptance correction for two τ_h passing the τ_h-selection, including the 45 GeV threshold, and tight isolation criteria, is determined. The probability of a trigger leg l of trigger t to fire is calculated by using the number of events passing the trigger N^{pass} and the number of events failing the trigger N^{fail}. For the case of the DiTau trigger used, two τ_h on the trigger efficiency plateau after selection are expected to have similar efficiencies. Therefore, the efficiency of both trigger legs with the same isolation is multiplied and the overall number of passing and failing events is the square of the efficiency of a single trigger leg, as described in Eq. 6.14.

$$P(l|\text{T}^{\text{real}}) = \sqrt{P(t|\text{T}_1^{\text{real}}, \text{T}_2^{\text{real}})} = \sqrt{\frac{N^{\text{pass}}}{N^{\text{pass}} + N^{\text{fail}}}} \tag{6.14}$$

On the W+jets samples, events with one real τ_h lepton and one τ_h^{fake} object are used to determine the trigger acceptance for looser isolation requirements for real τ_h leptons and for all three different exclusive τ_h^{fake} object isolation criteria. In principle, any sample with one real and one fake τ_h object is usable, but the W+jet samples have a sufficient amount of events, by themselves. The entries in Tab. 6.11 passed the HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Prong1_v* trigger and the entries in Tab. 6.12 did not pass the same trigger. All events are required to pass τ_h lepton selections, including the 45 GeV threshold.

	T^{real}	M^{real}	L^{real}
L^{fake}	252	28	22
M^{fake}	202	9	26
T^{fake}	472	51	40

TABLE 6.11: Table of rounded weighted event numbers passing the trigger out of W+jet simulation.

	T^{real}	M^{real}	L^{real}
L^{fake}	464	69	41
M^{fake}	151	26	15
T^{fake}	288	48	52

TABLE 6.12: Table of rounded weighted event numbers failing the trigger out of W+jet simulation.

To calculate the acceptance of the di-tau trigger for τ_h^{fake} objects, the efficiency of the trigger on the mixed real and fake τ_h events is assumed to be a multiplication of the efficiency of the real and the fake τ_h trigger legs. Determining the trigger efficiency on a mixed event, the fake τ_h trigger leg efficiency is calculated by correcting for the real τ_h trigger leg efficiency, as shown in Eq. 6.15.

$$P(l|\text{ID}^{\text{fake}}) = \frac{N^{\text{pass}}(\text{ID}^{\text{fake}}, T^{\text{real}})}{N^{\text{pass}}(\text{ID}^{\text{fake}}, T^{\text{real}}) + N^{\text{fail}}(\text{ID}^{\text{fake}}, T^{\text{real}})} \cdot \frac{1}{P(l|T^{\text{real}})} \qquad (6.15)$$

To calculate the acceptance of the trigger for looser real τ_h-lepton isolations, both trigger legs are assumed to be independent from each other, allowing the probability of a given ID^{real} trigger leg to fire the trigger in association with any $\text{ID}^{\text{fake, incl}} = \{T^{\text{fake}}, M^{\text{fake}}, L^{\text{fake}}\}$ to be derived in Eq. 6.16.

$$P(t|\text{ID}^{\text{real}}, \text{ID}^{\text{fake, incl}}) = \frac{N^{\text{pass}}(\text{ID}^{\text{fake, incl}}, \text{ID}^{\text{real}})}{N^{\text{pass}}(\text{ID}^{\text{fake, incl}}, \text{ID}^{\text{real}}) + N^{\text{fail}}(\text{ID}^{\text{fake, incl}}, \text{ID}^{\text{real}})} \qquad (6.16)$$

One can then derive the probability of a trigger leg registering a given τ_h^{fake} object of a certain isolation ID to contribute to firing the trigger in conjunction with a real τ_h lepton of another isolation ID, using Eq. 6.17.

$$P(l|\text{ID}^{\text{real}}) = \frac{P(t|\text{ID}^{\text{real}}, \text{ID}^{\text{fake, incl}})}{P(t|T^{\text{real}}, \text{ID}^{\text{fake, incl}})} \cdot P(l|T^{\text{real}}) \qquad (6.17)$$

The results of this simulated trigger study are in Tab. 6.13 with a gaussian uncertainty propagation uncertainty.

	$\epsilon^{\text{trigger}}(\text{ID})$	$\sigma(\epsilon^{\text{trigger}}(\text{ID}))$
T^{real}	0.836	± 0.009
M^{real}	0.376	± 0.007
L^{real}	0.330	± 0.008
T^{fake}	0.771	± 0.012
M^{fake}	0.641	± 0.008
L^{fake}	0.426	± 0.007

TABLE 6.13: Table of simulated trigger efficiencies for exclusive isolation criteria of real and fake τ_h objects with statistical uncertainties.

The main uncertainty of this method is, due to the statistical limitation by the samples available, the assumption that the $p_T^{\tau_h} \geq 45\,\text{GeV}$ threshold tested and determined on real τ_h passing tight isolation criteria is still valid for looser isolation requirements and/or τ_h^{fake} objects. It would be desirable, for future studies, to not trigger on the reconstructed τ_h leptons, in order to forego the reliance on a simulated trigger acceptance value, to not need such a high transverse momentum threshold and to more accurately model the $p_T^{\tau_h}$ shape without additional uncertainties due to missing knowledge of the trigger turn-on for looser isolation requirements.

6.2.6 Fixation of fake τ_h mass

Out of Ref. [49] one can deduce the proportionality of Eq. 6.18 for jets:

$$< M^2 >= constant \cdot p_T^2 \cdot R^2 \qquad (6.18)$$

The jet transverse momentum spectrum is an exponentially decreasing function. Therefore, by Eq. 6.18, the expected mass spectrum for jets is exponentially decreasing, as well. In essence, the jet mass spectrum is correlated with p_T.

Taking this into account and the properties of jet to τ_h momentum translation derived in Sec. 6.2.3, the reconstructed invariant mass shape of τ_h^{fake} should be similar to that of jets, but there should be an offset caused by the difference in effective cone size of the jet and the τ_h object. A calorimeter cell has a side length of $\Delta_R \approx 0.1$, a standard anti-kt 5 jet in the CMS experiment has a cone size of $\Delta R = 0.5$. With a constant translation factor for the transverse momentum of jet to τ_h, the difference caused by the cone size ratio, alone, would therefore be $\frac{\Delta R}{\Delta_R} \approx \mathcal{O}(5)$. The momentum translation factors plateau at roughly $p_T^{jet} \rightarrow p_T^\tau \approx 0.6$. Therefore, one can actually predict the ratio of single jet

mass to single fake τ_h mass, as done in Eq. 6.19.

$$\frac{M^{\tau_h^{\text{fake}}}}{M^{\text{jet}}} = \frac{0.6 \cdot p_T^{\text{jet}} \cdot 0.1}{1 \cdot p_T^{\text{jet}} \cdot 0.5} = 0.12 \qquad (6.19)$$

According to this estimate, one can expect the reconstructed τ_h^{fake} mass to always be smaller than the jet mass by an order of magnitude. For this reason, one cannot simply set the diced τ_h^{fake} mass to the jet mass. The distribution of identified τ_h masses and their corresponding jet (matched within $\Delta R < 0.1$) is shown in Fig. 6.16 for various isolations out of simulated QCD samples.
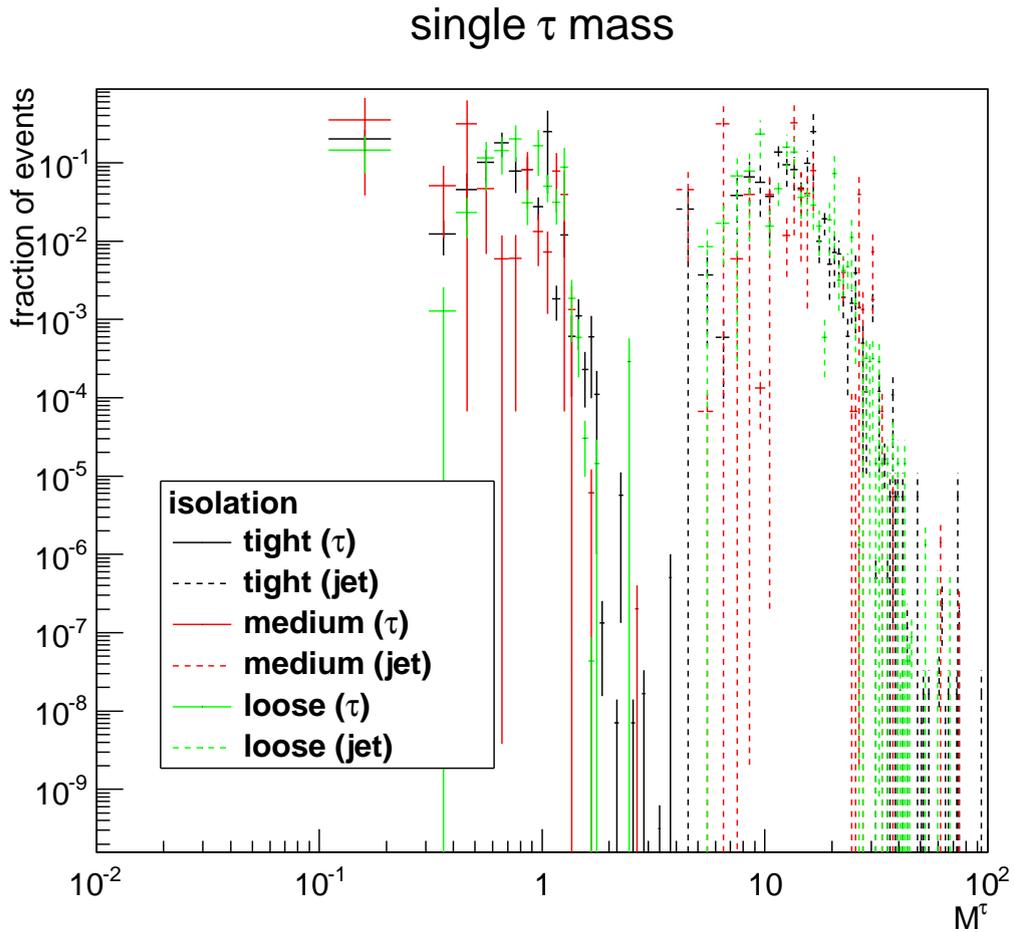
FIGURE 6.16: Reconstructed invariant mass of the τ_h objects and their corresponding jet objects matched within $\Delta R < 0.1$. It can be observed that the reconstructed mass is independent of isolation and that there is a constant shift of the spectrum.

The predicted trend can be verified and the average of the τ_h^{fake} mass has been determined in data to be 720 MeV. All jets transformed to τ_h^{fake} are set to this mass. The

difference of the shape of the invariant mass distribution of the di-τ_h^{fake} system with and without this correction, as well as a partial correction only taking into account the jet to τ_h transverse momentum translation factors is shown for QCD events in Fig. 6.17.
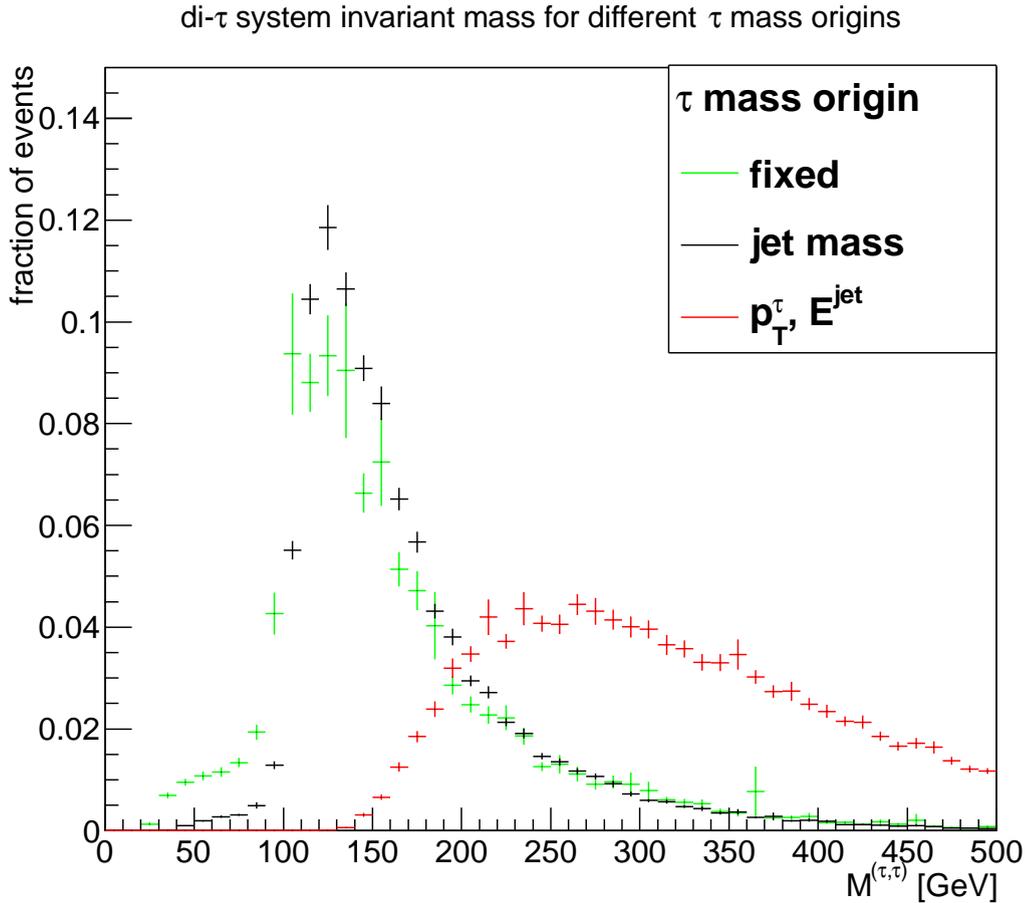
FIGURE 6.17: Shape of the reconstructed invariant mass of the di-τ_h^{fake} system, shown with (green) and without (black) applying a reasonable single τ_h^{fake} mass value. Another distribution is shown, correcting the transverse momentum differences, but not attributing for any changes in overall energy (red). Such a partial correction is a grave mistake, as it shifts the mass of the di-τ_h system to large masses.

In order to show the feasibility of the mass reconstruction with this correction, the invariant mass values of data and backgrounds in opposite sign control region two (two τ_h passing tight isolation criteria) is shown in Fig. 6.18. With the Z-boson resonance as a reference, one can observe a reasonable description of the mass distribution for a purely simulation-driven method.
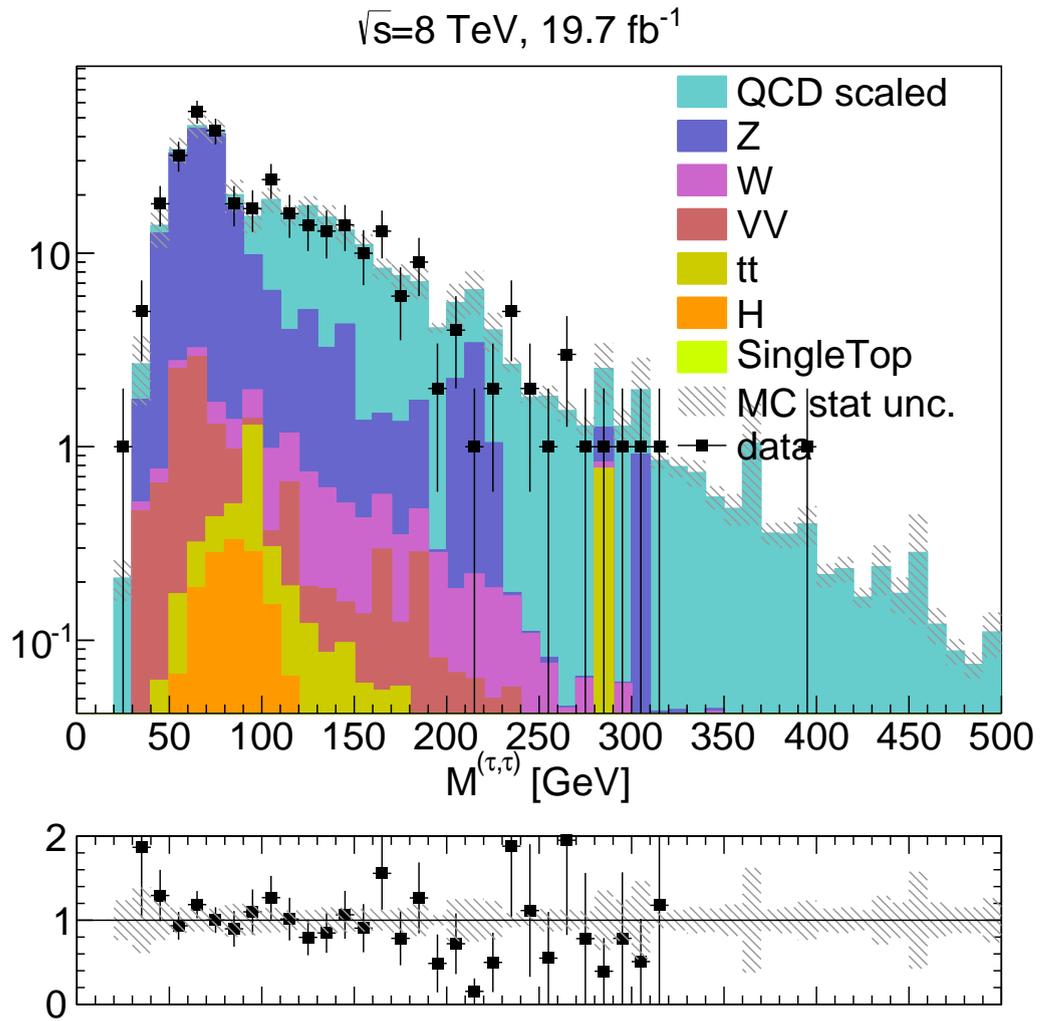
FIGURE 6.18: Comparison of reconstructed invariant mass of the di-τ_h system in the opposite sign control region two with two τ_h passing tight isolation criteria. The single τ_h mass correction is applied for all backgrounds with diced events.

6.3 Results

The main improvement by using dicing techniques on backgrounds is best summarized by Tab. 6.14.

control	OS		LS	
	nominal	diced	nominal	diced
SR	0	$5 \cdot 10^5$	0	$4 \cdot 10^5$
CR 2	5	$5 \cdot 10^6$	4	$3 \cdot 10^6$
CR 3	0	10^6	0	$7 \cdot 10^5$
CR 4	10	$9 \cdot 10^6$	18	$6 \cdot 10^6$
CR 5	0	10^6	0	$7 \cdot 10^5$
CR 6	4	$9 \cdot 10^6$	8	$6 \cdot 10^6$
CR 7	0	10^6	0	$7 \cdot 10^5$
CR 8	4	$9 \cdot 10^6$	6	$6 \cdot 10^6$

TABLE 6.14: Table of the number of simulated events for QCD on which any prediction for any region is based on, depending on whether the dicing method is applied (diced) or not applied (nominal).

Despite a total sample size of 10^8, not a single event with two fake τ_h passing VBF and central selections has been observed, before the application of dicing. Thus, it is without any realistic alternative for having any prediction based on simulation to use a dicing method.

In the following, the descriptive power of the method in terms of shape comparisons to data are shown in Sec. 6.3.1, systematic uncertainties are determined in Sec. 6.3.2 and the exclusion limits on vector boson fusion supersymmetry models by the VBF search are shown in Sec. 6.3.3. Finally, possible future improvements on the background method will be outlined in Sec. 6.3.4.

6.3.1 Descriptive power of diced simulation

In this section, examples of data to diced simulation comparison are shown in order to point out the achievements and failures of diced simulation, so far. All shown QCD distributions are scaled in order to exactly match the difference between the amount of data measured to the amount of non-QCD background events expected in the control region with the exception of the signal region where the scaling is done to the prediction of the ABCD method performed on data. A complete set of all distributions in linear and logarithmic versions are attached in App. B.4.

First, one can look at the overall background composition in different control regions. There are major differences, depending on whether the region is opposite sign or like sign and which isolation requirements need to be fulfilled.

(a) OS CR2 (TT)

(b) LS CR2 (TT)

(c) LS CR4 (TM$_I$)
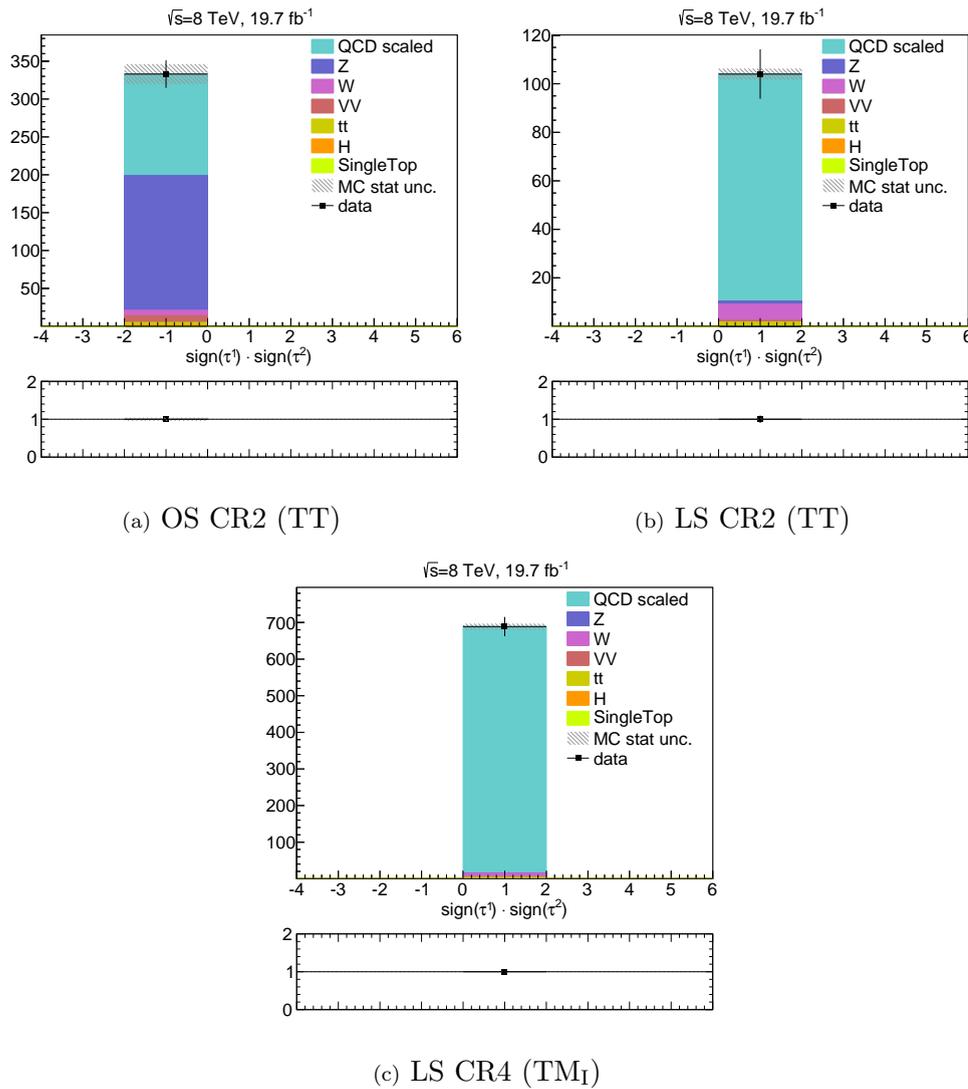
FIGURE 6.19: Background composition for different control regions ranging from a large contamination in the opposite sign region to $\approx 10\%$ contamination in its like sign counterpart to basically pure QCD when requiring looser isolation on like sign. QCD is scaled to the number of events in data minus non-QCD background contributions.

In general, the opposite sign regions like Fig. 6.19(a) are far more contaminated with non-QCD backgrounds than their like sign counterparts like Fig. 6.19(b) and any background producing at least one real τ_h is a lot more likely to be in a control region requiring at least one tightly isolated τ_h. Therefore, requiring more loosely isolated τ_h increases QCD purity, as seen in Fig. 6.19. Still, there are other distributions enabling one to

differentiate between QCD and other backgrounds like Drell-Yang processes. One such distribution is the angle between the two τ_h, shown in Fig. 6.20

(a) OS CR2 (TT) (b) LS CR4 (TM$_I$)
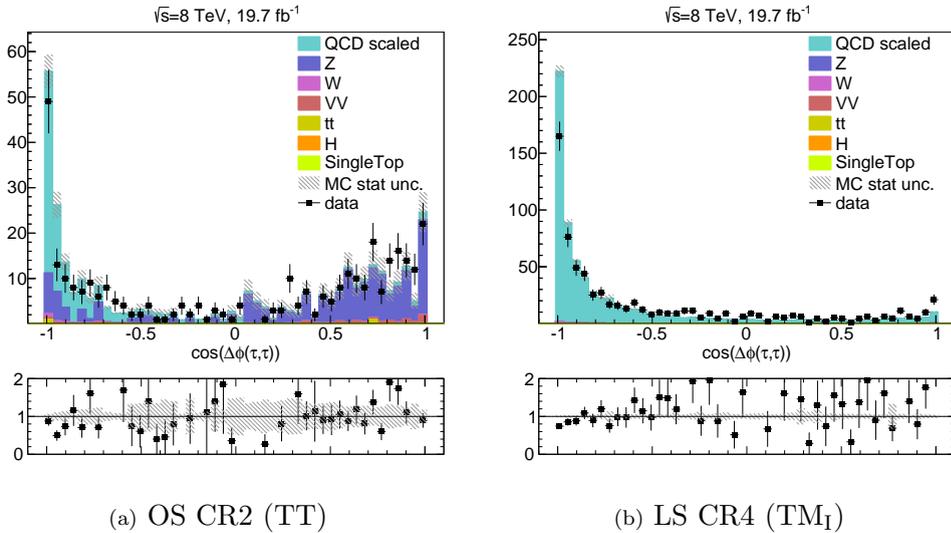
FIGURE 6.20: Cosine of the separation angle in ϕ of the two τ_h in the event. For QCD, back to back topologies are the most likely to occur, so the leading two jets are most likely to be chosen. On the other hand, Drell-Yang processes feature small separation angles as the decay products of the Z boson are boosted.

It can be seen that the description is not perfect and that there are more events with a small separation of the τ_h which can either be explained by an unmodeled contribution of charge-misreconstructed Drell-Yang events or by choosing the wrong jets to become τ_h in some cases. The latter would require more low momentum jets out of initial or final state radiation or pileup to be chosen, because the angle would be random and broadly enhance the $\cos(\Delta\phi(\tau_1\tau_2))$, which in turn would soften the transverse momentum spectrum. Such a softening would improve the agreement of the diced simulation to data, as can be seen when looking at the τ_h transverse momenta in Fig. 6.21. As shown in Fig. 6.21(b), the softer part of the event is not described. This fits to the observations made in Sec. 6.2.4 that the transmutation factors do not work at low transverse momenta, correctly.

It is not possible to locate a specific detector area in which this difference originates, as shown in Fig. 6.22. This is expected, because the transmutation factors are the the main cause of the low transverse momentum underprediction and they do not depend on $|\eta|$.

Observables of the di-τ_h system are better described, as shown in Fig. 6.23(b) for the separation of both τ_h in η and in Fig. 6.23(c) or 6.23(d) for the invariant mass of both objects.

(a) Leading τ_h in LS CR4 (TM$_I$)

(b) Second to leading τ_h in LS CR4 (TM$_I$)
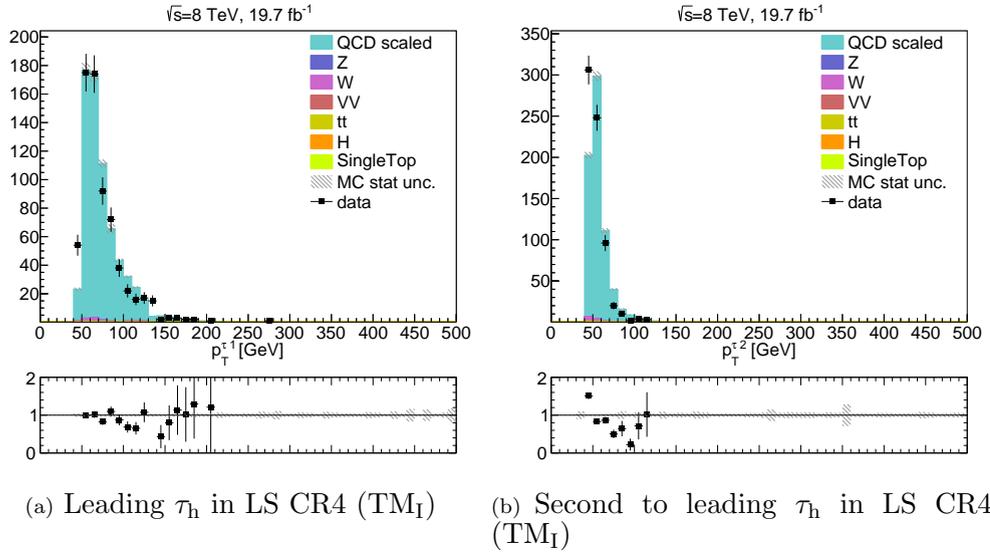
FIGURE 6.21: Transverse momentum of the leading and second to leading τ_h. A good agreement can be observed for the leading τ_h, safe for the first bin. On the other hand, the second to leading τ_h momentum is badly described.

(a) Leading τ_h in LS CR4 (TM$_I$)

(b) Second to leading τ_h in LS CR4 (TT)
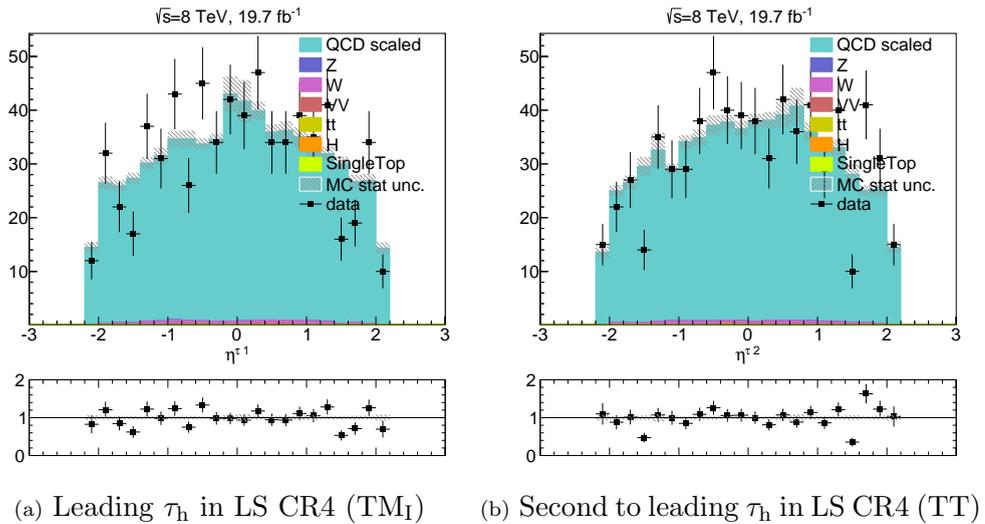
FIGURE 6.22: Pseudorapidity of the leading and second to leading τ_h. A substantial difference can be seen at $|\eta| \approx 1.3$, because the τ ID vetos the barrel to endcaps tracker transition region. This is not taken into account by the dicing method at all, as the jet to τ_h axis is unstable within $\Delta R \leq 0.1$ making a similar veto problematic while the overall difference is small.

(a) OS CR2 (TM$_I$)

(b) LS CR4 (TM$_I$)

(c) OS CR2 (TT)

(d) LS CR4 (TM$_I$)
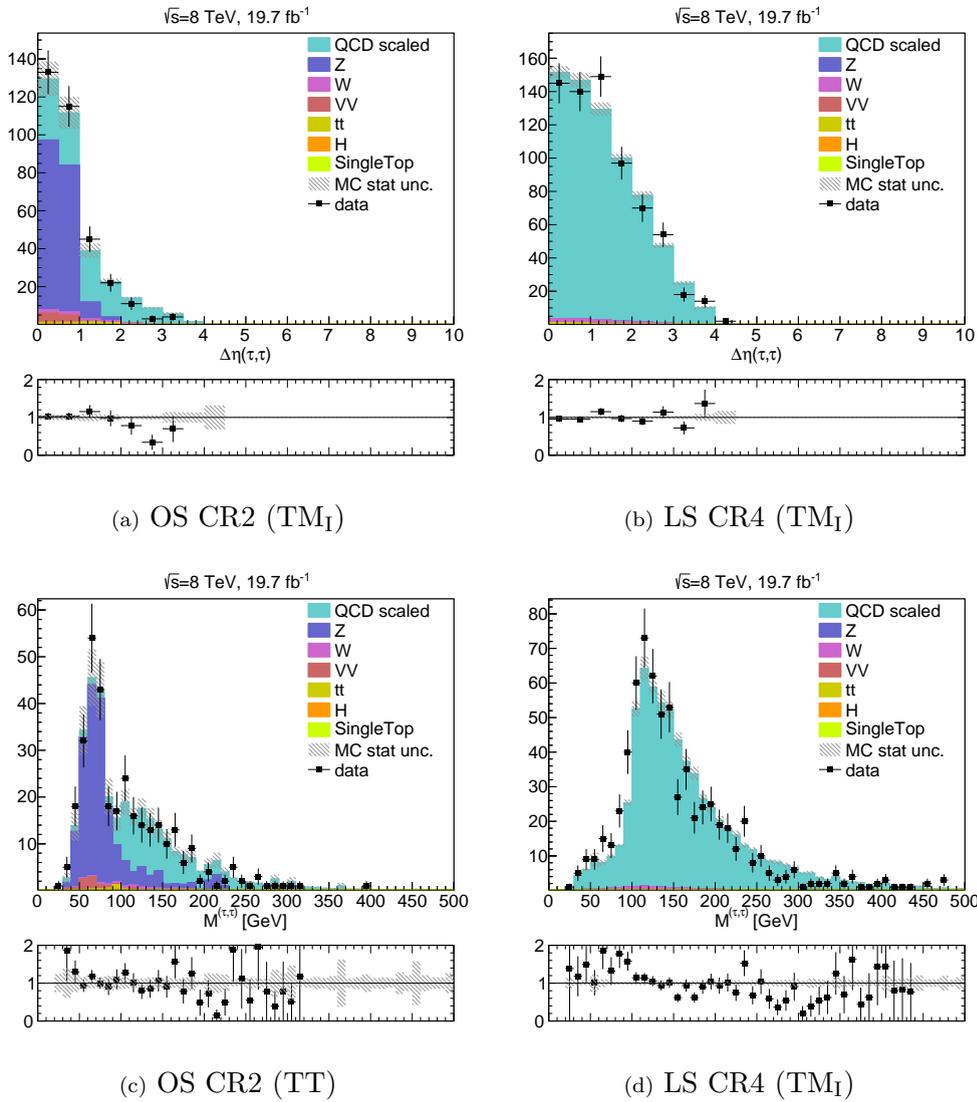
FIGURE 6.23: Shown are a/b) $\Delta\eta(\tau_1, \tau_2)$ and c/d) $M(\tau_1, \tau_2)$. There is good agreement between data and diced simulation that is remarkable considering the disagreement regarding the second to leading τ_h transverse momentum.

Another kind of observable are jet based observables. The sum of all transverse momenta of all jets passing selection requirements is called H_T and shown in Fig. 6.24.

(a) LS CR4 (TM$_I$)   (b) OS CR8 (LL)
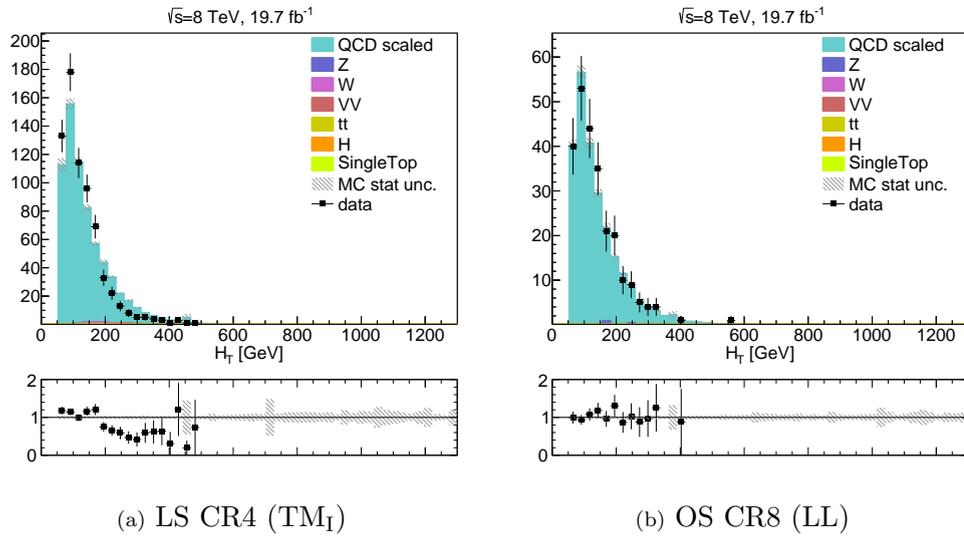
FIGURE 6.24: Sum of all jet transverse momenta. Not all isolations are equally well described. Requiring at least one τ_h passing tight isolation criteria leads to an overprediction of events with high-energetic jets.

In order to determine which jets are resonsible for the mismatch observed, Fig. 6.25 shows the distribution of leading, second to leading and all jet transverse momenta.
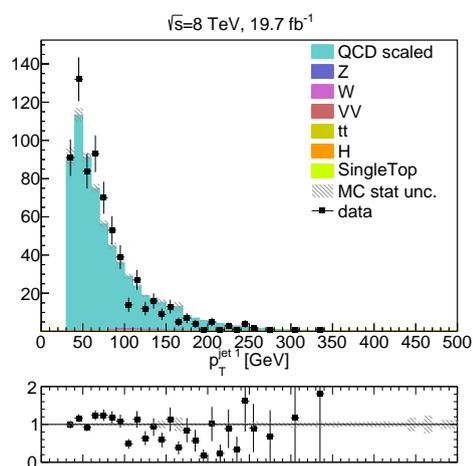
(a) leading jet in LS CR4 (TM$_I$)
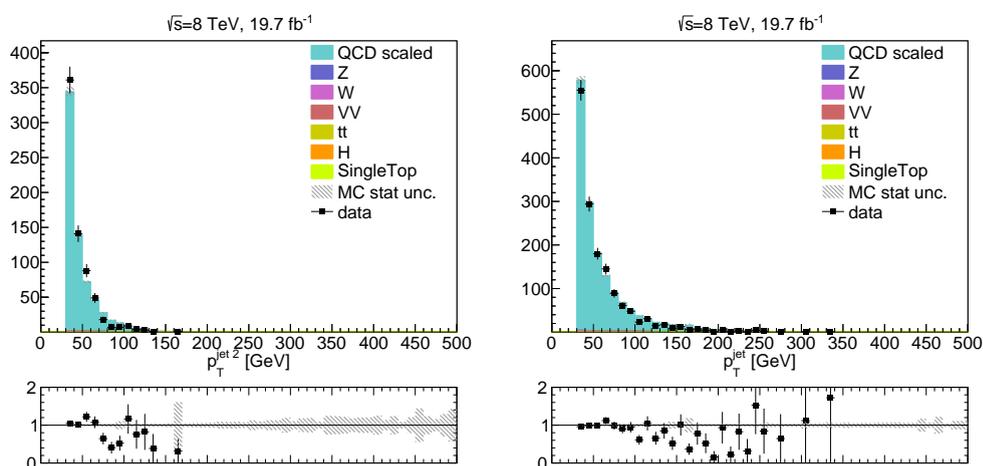
(b) second to leading jet in LS CR4 (TM$_I$)

(c) all jets in LS CR4 (TM$_I$)

FIGURE 6.25: Transverse momentum of the a) leading jet b) second to leading jet and c) all jets. The quality of the description of the distributions deteriorates in the high transverse momenta tails.

Unlike for the τ_h, a source of the discrepancy can be localized to the $2 < |\eta| < 3$ region of the detector, as shown in Fig. 6.26.
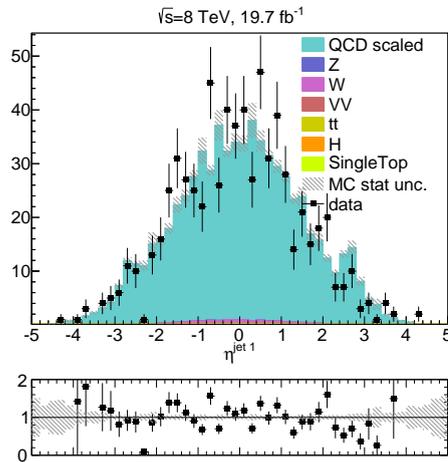
(a) leading jet in LS CR4 (TM$_I$)
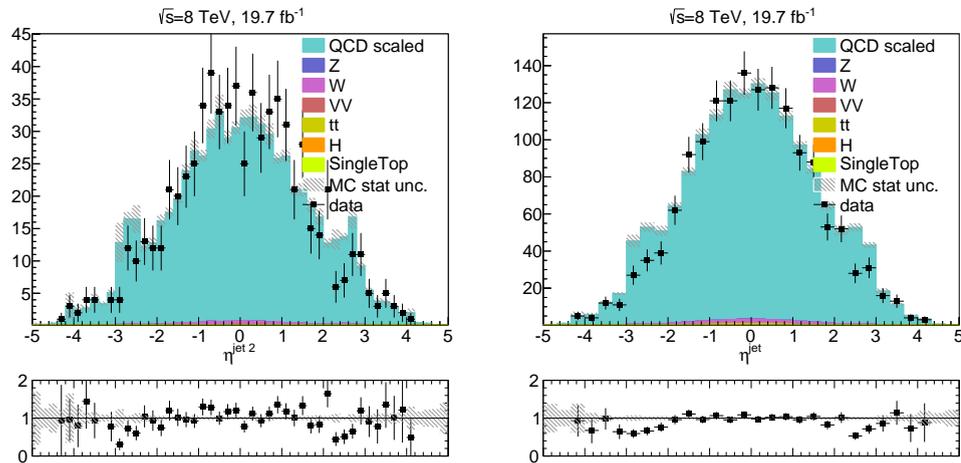
(b) second to leading jet in LS CR4 (TM$_I$)

(c) all jets in LS CR4 (TM$_I$)

FIGURE 6.26: Pseudorapidity of the a) leading jet b) second to leading jet and c) all jets. There is an overprediction in the endcap region that is observable for the second to leading jets and especially for all jets, inclusively. For the leading jets, such an observation is not conclusive.

Furthermore, the spectrum of the number of jets disagrees, as seen in Fig. 6.27.

(a) OS CR2 (TT) (b) LS CR4 (TM$_{\mathrm{I}}$)
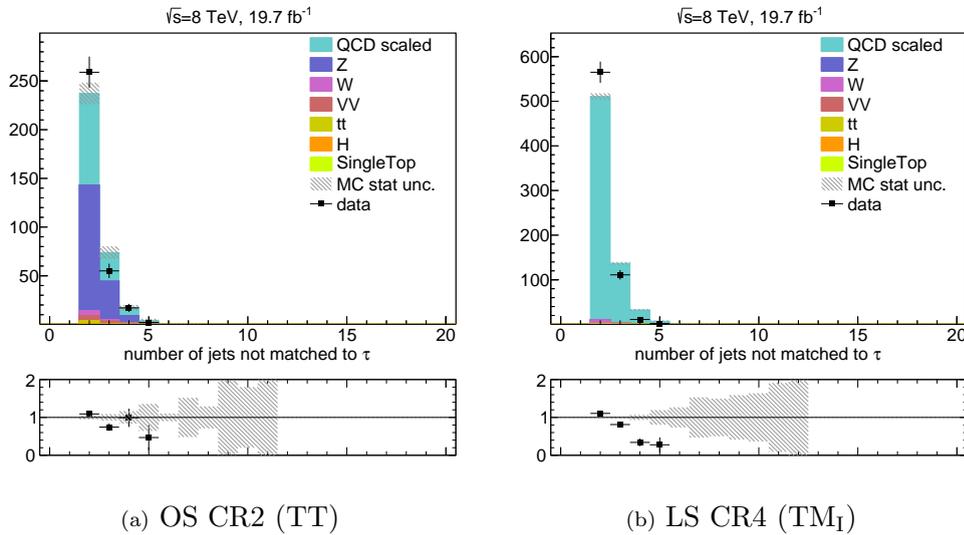
FIGURE 6.27: Number of jets passing all selection requirements. The leading order contribution fits by default due to the scaling of the QCD to the number of events in data minus simulated non-QCD contributions. A nonclosure can be observed for the QCD contribution. Other standard model backgrounds seem to be described better.

As samples generated with PYTHIA were used, using MADGRAPH instead might be a possible improvement, because MADGRAPH would be expected to better model higher order contributions to the matrix element than PYTHIA by itself.

Finally, the missing transverse momentum is shown in Fig. 6.28 for a VBF-like control region and its inversion, as well as the pseudorapidity distribution of all jets for these two regions.
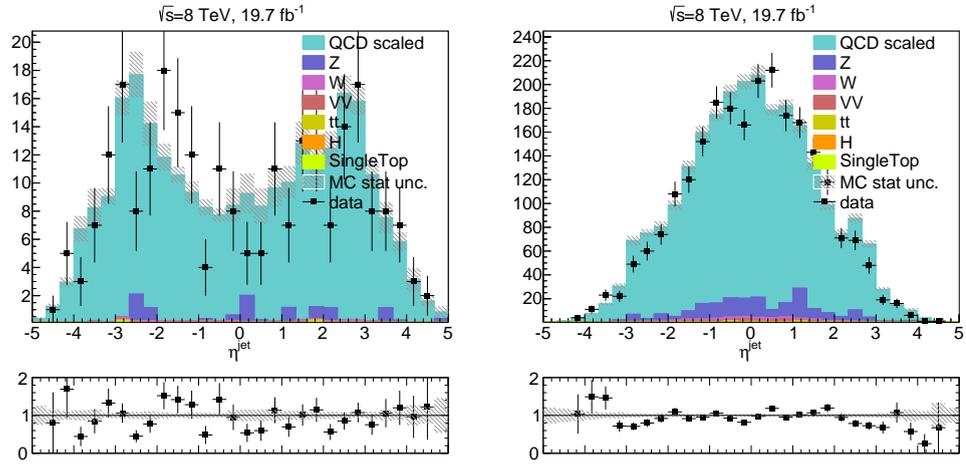
(a) Pseudorapidity of all jets in OS CR3 (TM$_I$)

(b) Pseudorapidity of all jets in OS CR4 (TM$_I$)
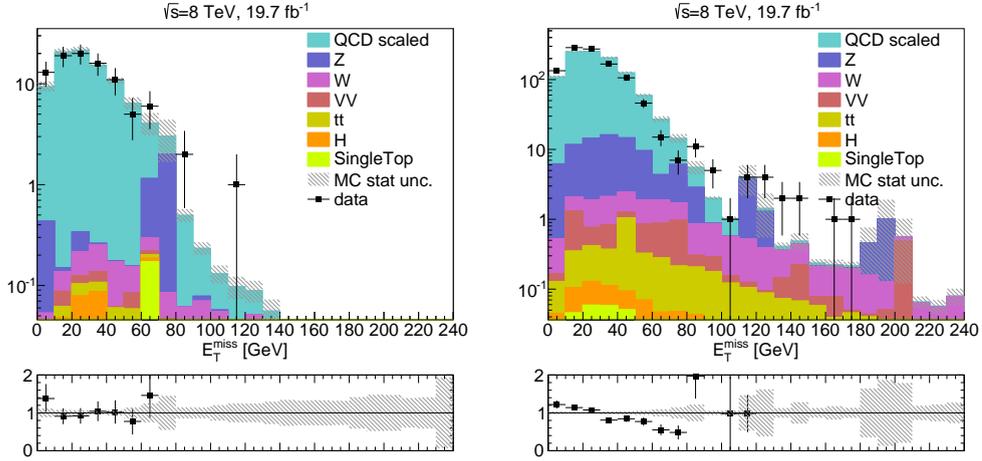
(c) $\not{E}_T$ in OS CR3 (TM$_I$)

(d) $\not{E}_T$ in OS CR4 (TM$_I$)

FIGURE 6.28: a/b) Pseudorapidity of all jets in a) VBF and b) VBF inverted regions requiring one tight and one medium or loose τ_h. c/d) $\not{E}_T$ in c) VBF and d) VBF inverted regions requiring one tight and one medium or loose τ_h.

A clear selection bias of the jet pseudorapidity can be observed by comparing Fig. 6.28(a) to Fig. 6.28(b). This leads to jets with different resolutions (see Fig. 5.16(a) for JER smearing factors with respect to $|\eta|$) being selected.

Moreover, in a QCD event there is no expectation of real missing transverse momentum. Therefore, the measured $\not{E}_T$ stems from mismeasurements of the jets. The broader the resolution, the larger the average $\not{E}_T$ per event. The resolution in the detector is worst in the endcaps and hadronic forward parts of the detector, thus biasing the events passing VBF selections towards higher average $\not{E}_T$.

Lastly, Tab. 6.15 shows all scaling factors used for all control regions and the signal

region. It can be observed that opposite sign regions have more data entries than like

LS isolation	VBF	$\overline{\text{VBF}}$
T + T	1.04	2.23
T + M/L	0.80	2.14
M + M/L	1.29	1.96
L + L	0.97	1.71
OS isolation	VBF	$\overline{\text{VBF}}$
T + T	2.19	3.27
T + M/L	2.23	3.00
M + M/L	1.75	2.44
L + L	1.30	2.29

TABLE 6.15: Table of scale factors needed to match diced QCD to data, T is tight isolation, M is medium, L is loose

sign regions (in unscaled simulation, there is no difference between signs) and that the VBF efficiency is mismodeled by a factor of 2. Still, within one sign and one VBF category, the spread of the scale factors is tolerable.

6.3.2 Systematic uncertainties determined on diced simulation

The efficiency of the VBF selections from Eq. 6.2 is the paramount subject to systematic uncertainties in the background estimation. In this section, a general approach to defining a VBF efficiency is pursued in order to determine which systematic uncertainties may occur and how to appraise them, in the end.

In a simplified form, the basic assumption made for the background estimation to be valid is that the measurement of ϵ^{VBF} is independent of the τ_{h} isolation. The most generic definition of a VBF efficiency is to state that the amount of events with at least four jets out of which two are misreconstructed as $\tau_{\text{h}}^{\text{fake}}$ passing central selection criteria and two other jets satisfying VBF selection criteria are defined as passing events $N_{4\,\text{jets}}^{\text{VBF}}$. On the other hand, events with at least two jets misreconstructed as $\tau_{\text{h}}^{\text{fake}}$ passing central selection criteria but failing VBF selection criteria are defined as failing events $N_{2+\text{n jets}}^{\overline{\text{VBF}}}$. In principle, more jets could be required for the definition of the failing region. This is allowed in the nomenclature by specifying n as a free parameter for additional jets. Furthermore, the exclusive isolation of the $\tau_{\text{h}}^{\text{fake}}$ must be specified. Here, this will be done by explicitly stating $N_{4\,\text{jets}}^{\text{VBF}}(\text{ID},\text{ID})$. In such a case, the efficiency as defined in Eq. 6.20 in which T stands for tight and M_{I} for medium or loose isolation must hold true.

$$\frac{N_{4\,\text{jets}}^{\text{VBF}}(\text{T},\text{T})}{N_{4\,\text{jets}}^{\text{VBF}}(\text{T},\text{T}) + N_{2+\text{n jets}}^{\overline{\text{VBF}}}(\text{T},\text{T})} = \frac{N_{4\,\text{jets}}^{\text{VBF}}(\text{T},M_{\text{I}})}{N_{4\,\text{jets}}^{\text{VBF}}(\text{T},M_{\text{I}}) + N_{2+\text{n jets}}^{\overline{\text{VBF}}}(\text{T},M_{\text{I}})} \quad (6.20)$$

Thus, the ratio of events passing VBF selections to all events passing and failing VBF selections should be the same no matter whether the two τ_h required pass tight isolation criteria or one passes tight and the other passes medium or loose isolation criteria.

Any event selection like $N_{2+n\text{ jets}}^{\text{VBF}}(\text{T},\text{M}_\text{I})$ is proportional to the fake probabilities $P(\text{T})$ determining the expected abundance of τ_h^{fake} of the specified isolation (described in more detail in Eq. 6.7) and the selection efficiency of the VBF selections ϵ^{VBF} which ought to be measured, as described in Eq. 6.21 and for events failing VBF selections in Eq. 6.22.

$$N_{4\text{ jets}}^{\text{VBF}}(\text{T},\text{M}_\text{I}) = \epsilon^{\text{VBF}} N_{4\text{ jets}} P(\text{T}) P(\text{M}_\text{I}) (1 - P(\text{T}_\text{I}))^2 \tag{6.21}$$

$$N_{2+n\text{ jets}}^{\overline{\text{VBF}}}(\text{T},\text{M}_\text{I}) = \epsilon^{\overline{\text{VBF}}} N_{2+n\text{ jets}} P(\text{T}) P(\text{M}_\text{I}) (1 - P(\text{T}_\text{I}))^n \tag{6.22}$$

The number of jets expected in a QCD event is exponentially falling as each jet more requires one more order of α_s and the transverse momentum spectrum is also different, because the distribution of energy among the correlated jets depends on the number of jets in the event. Furthermore, each reconstructed or misreconstructed jet needs at least 30 GeV of transverse momentum, in order to pass selection requirements. Therefore, choosing $n = 0$ has the advantage of superior statistics in the denominator of any side of Eq. 6.20. A disadvantage might be a correlation of jet transverse momentum to the VBF efficiency. Should such a correlation exist, having a different jet transverse momentum spectrum for passing and failing event definitions would lead to a systematic bias of the VBF efficiency definition.

In order to demonstrate how such a bias could occur, a jet transverse momentum dependence of the VBF efficiency has to be deducted, as well as a correlation in the definition of the VBF efficiency itself. To this end, a first step is to quantify the impact of the choice of n without introducing p_T dependencies. Thus, each side of Eq. 6.20 can be written expressly using Eq. 6.21 and 6.22, as demonstrated in Eq. 6.23.

$$\frac{\epsilon^{\text{VBF}} N_{4\text{ jets}} P(\text{T}) P(\text{M}_\text{I}) (1 - P(\text{T}_\text{I}))^2}{\epsilon^{\text{VBF}} N_{4\text{ jets}} P(\text{T}) P(\text{M}_\text{I}) (1 - P(\text{T}_\text{I}))^2 + \epsilon^{\overline{\text{VBF}}} N_{2+n\text{ jets}} P(\text{T}) P(\text{M}_\text{I}) (1 - P(\text{T}_\text{I}))^n} \tag{6.23}$$

This equation can be simplified further to Eq. 6.24.

$$\left[1 + \frac{1 - \epsilon^{\text{VBF}}}{\epsilon^{\text{VBF}}} \frac{N_{2+n\text{ jets}}}{N_{4\text{ jets}}} (1 - P(\text{T}_\text{I}))^{n-2}\right]^{-1} \tag{6.24}$$

For $n = 2$, each side of Eq. 6.20 and therefore Eq. 6.24 perfectly evens out to ϵ^{VBF} and thus the very quantity that one wants to measure. Requiring less jets introduces a small bias, because the small fake probabilities lead to a factor of $(1 - P(\text{T}_\text{I}))^{n-2} \approx 1$.

Still, this picture is too simple as it depends on the assumption that ϵ^{VBF} is uncorrelated to p_T and therefore the fake probabilities. That can be tested by deriving the VBF

efficiency differentially as a function of, e.g., $p_T^{\tau_h^{\text{fake}}}$ which is by the transmutation factors a direct function of p_T, as done in Fig. 6.29.
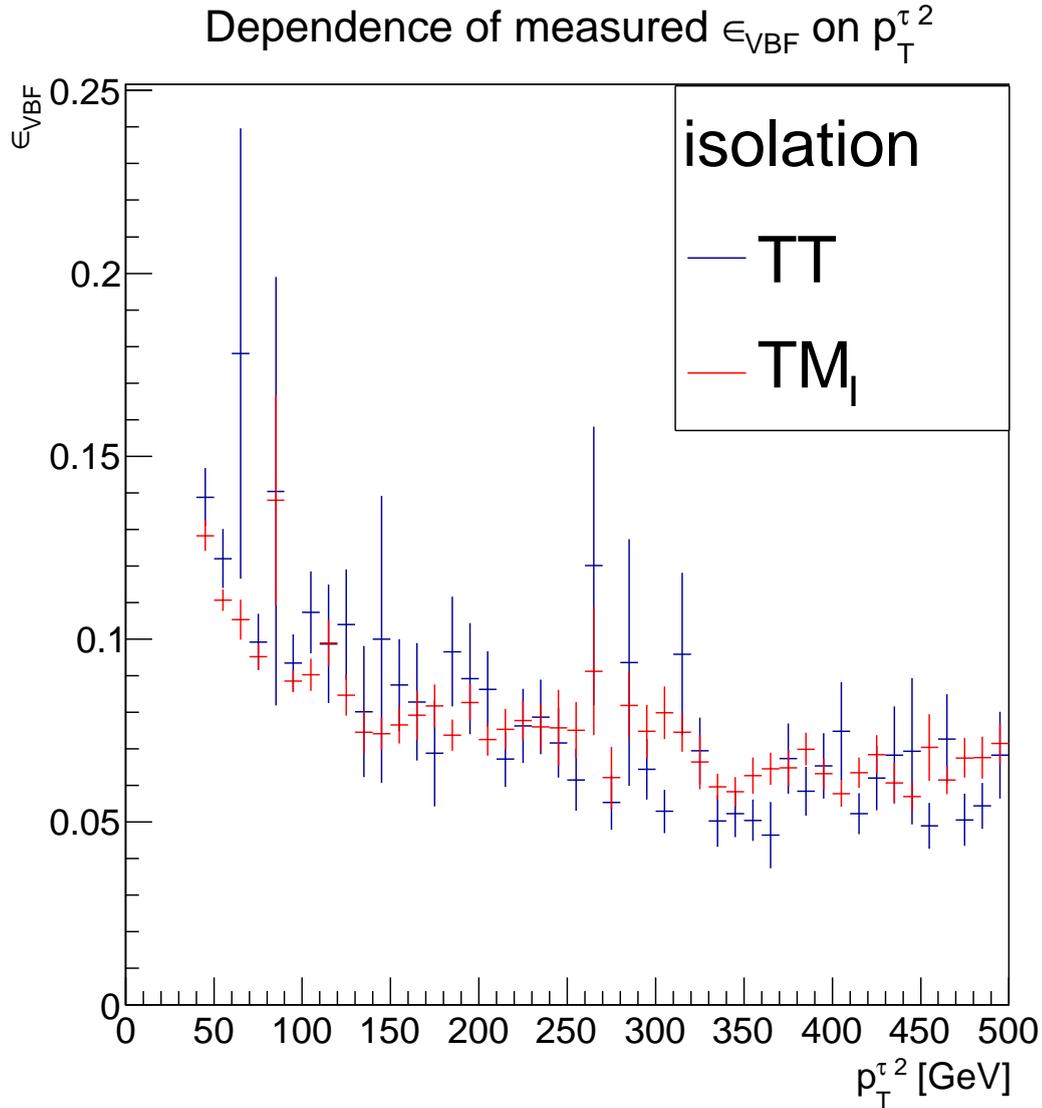
FIGURE 6.29: Measurement of ϵ_{VBF} by bin of second to leading τ_h momentum for the requirement of two τ_h passing tight (TT) or lower (TM$_{\text{I}}$) isolation criteria in diced QCD simulated events. A turn on can be observed, before the measurement converges against a flat ϵ_{VBF} at high momenta.

Thus, the turn on of the efficiency measurement is independent of isolation, but ϵ^{VBF} is not determined differentially on data. Instead, the number of events is counted. This corresponds to a sum over conditional probabilities of the four jets i, j, k, l per event as shown in Eq. 6.25. Each condition implicitly contains jet transverse momentum, because

$$P(\mathrm{T}, p_T^{\tau_h^{\text{fake}}}) \propto P(\mathrm{T}, p_T^{\text{jet}}).$$

$$N_{4\,\text{jets}}^{\text{VBF}}(\mathrm{T},\mathrm{T}) = \sum_{N_{4\,\text{jets}}} \left(\sum_{\text{jet } i,j,k,l} \epsilon^{\text{VBF}}(p_T^{\tau_h^{\text{fake}}}|i,j) \cdot P(\mathrm{T}|k) \cdot P(\mathrm{T}|l) \cdot (1 - P(\mathrm{T}|i)) \cdot (1 - P(\mathrm{T}|j)) \right)$$

$$(6.25)$$

To conclude, the assumption that ϵ^{VBF} and isolation are uncorrelated is only true, if the shape of all different isolations with respect to p_T^{jet} is similar and the leading order jet spectrum out of which fake objects are selected is the same for the VBF selection and its inversion. The former assumption is wrong as shown in Fig. 6.30, the latter may be correct, depending on the definition of the VBF inversion. For $n = 2$, possible biases are minimized. Therefore, two additional jets were required in addition to two τ_h^{fake}.
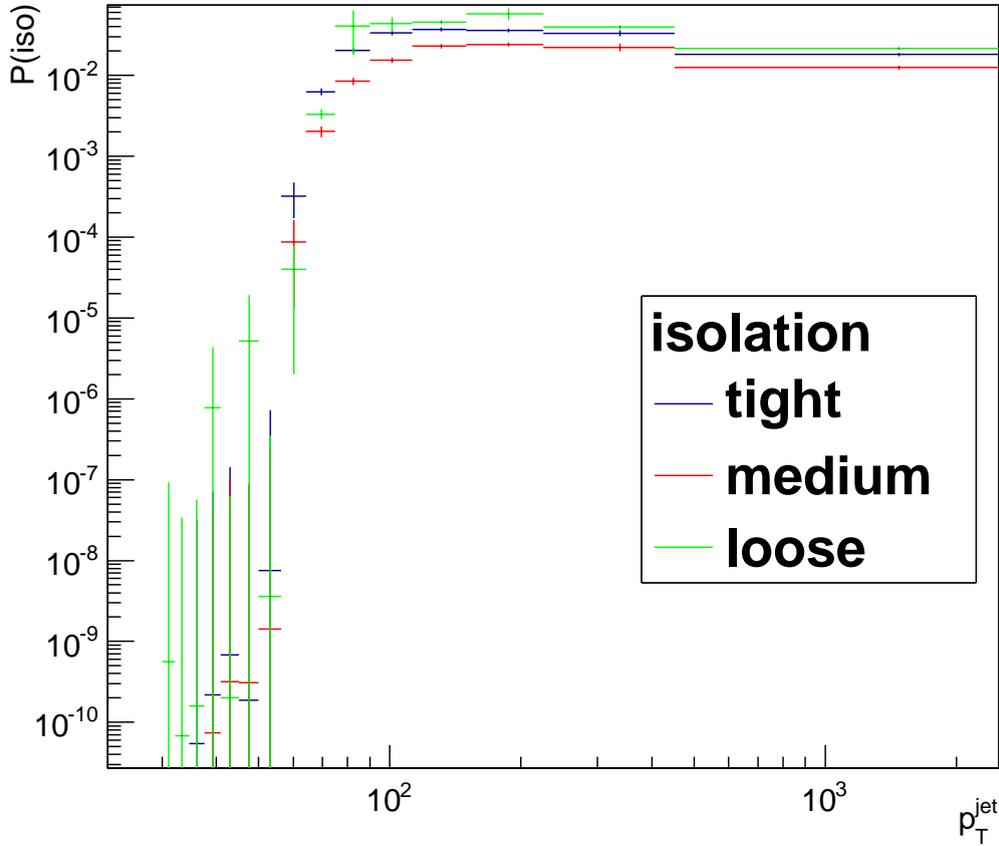
FIGURE 6.30: Projection of the fake probability maps out of Sec. 6.2.4 with respect to p_T^{jet} times the selection acceptance correction out of Sec. 6.2.2.

At least in simulation, distinct differences are observed. Thus, a systematic uncertainty has to be assigned for the level of disagreement expected in diced simulation, although this disagreement might be different in data. In order to test the stability of VBF

selections with respect to varying requirements on varying τ_h isolation selections and $\not{E}_T$, all values of ϵ^{VBF} in all even and odd control region pairs for a range of $\not{E}_T$-requirements are determined on the diced simulation and shown in Tab. 6.16.

region	$\epsilon^{\text{VBF}}(\not{E}_T \geq 0 \text{ GeV})$ [%]	$\epsilon^{\text{VBF}}(\not{E}_T \geq 10 \text{ GeV})$ [%]	$\epsilon^{\text{VBF}}(\not{E}_T \geq 20 \text{ GeV})$ [%]
LS SR/CR2	12.78 ± 0.53	12.91 ± 0.58	13.06 ± 0.67
LS CR3/CR4	12.20 ± 0.43	12.48 ± 0.47	12.46 ± 0.53
LS CR5/CR6	11.61 ± 0.41	11.93 ± 0.45	11.91 ± 0.50
LS CR7/CR8	12.10 ± 0.60	12.44 ± 0.66	12.75 ± 0.83
OS SR/CR2	11.46 ± 0.72	11.74 ± 0.80	12.23 ± 1.04
OS CR3/CR4	10.62 ± 0.31	10.66 ± 0.33	10.85 ± 0.36
OS CR5/CR6	10.98 ± 0.48	11.09 ± 0.52	11.59 ± 0.67
OS CR7/CR8	10.67 ± 0.36	10.61 ± 0.37	11.12 ± 0.47
WAM	11.33 ± 0.15	11.43 ± 0.16	11.66 ± 0.19
region	$\epsilon^{\text{VBF}}(\not{E}_T \geq 30 \text{ GeV})$ [%]	$\epsilon^{\text{VBF}}(\not{E}_T \geq 40 \text{ GeV})$ [%]	$\epsilon^{\text{VBF}}(\not{E}_T \geq 50 \text{ GeV})$ [%]
LS SR/CR2	13.76 ± 0.89	14.99 ± 1.49	18.09 ± 2.93
LS CR3/CR4	13.18 ± 0.68	14.54 ± 1.11	16.58 ± 2.08
LS CR5/CR6	12.20 ± 0.56	13.06 ± 0.84	13.34 ± 1.12
LS CR7/CR8	14.05 ± 1.25	15.93 ± 2.19	19.26 ± 4.14
OS SR/CR2	13.14 ± 1.58	16.04 ± 2.98	21.23 ± 6.13
OS CR3/CR4	11.08 ± 0.41	12.61 ± 0.69	14.45 ± 1.23
OS CR5/CR6	12.19 ± 0.95	14.38 ± 1.70	18.22 ± 3.42
OS CR7/CR8	11.37 ± 0.54	13.01 ± 0.90	14.76 ± 1.62
WAM	11.98 ± 0.23	13.43 ± 0.39	14.85 ± 0.65

TABLE 6.16: Stability of ϵ^{VBF} with respect to $\not{E}_T$, sign and τ_h-isolation. LS and OS regions are slightly correlated. Here, events can be the same, but the chosen jets to fake τ_h must differ. For efficiencies of one sign, efficiencies are expected to be highly correlated. This is not accounted for in the weighted arithmetic mean (WAM) values shown.

A quadratic dependence of ϵ^{VBF} on requirements on $\not{E}_T$ is observed, as seen in Fig. 6.31. This effect can be explained by the impact of higher jet energy resolution widths at more forward parts of the detector. According to Ref. [80], the expected average mass $\langle M \rangle$ reconstructed from two jets of similar energy E with a resolution σ_E and with an opening angle $a = 2(1 - \cos\theta)$ between them is $\langle M(E) \rangle \approx \sqrt{a}\langle E \rangle - \frac{\sqrt{a}\sigma_E^2}{8\langle E \rangle}$. Thus, the jet energy resolution and with it the occurence of $\not{E}_T$ in QCD events is directly tied to the mass of the di-jet system.
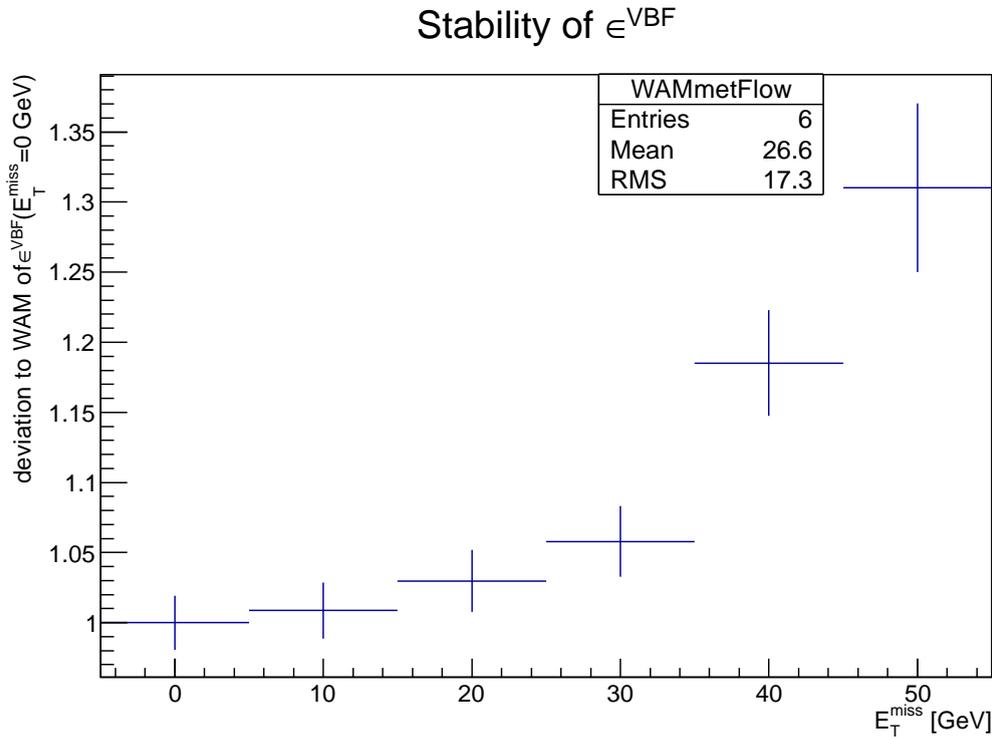
FIGURE 6.31: Deviation of the weighted arithmetic mean (WAM) of the VBF efficiency of all (LS and OS) control region ratios with respect to $\not{E}_\mathrm{T}$. A quadratic dependence of small size at the requirement of 30 GeV used in this analysis is observed and taken into account as a systematic uncertainty.

As seen in Sec. 6.2.3, the second to leading τ_h transverse momentum is ill described, such that at least estimate the size of the impact of different jet kinematics per τ_h isolation, the relative difference of the weighted arithmetic mean to the maximum and minimum difference out of Tab. 6.16 is taken into account as a systematic uncertainty on the stability with respect to τ_h-isolation. To summarize, two systematic uncertainties are taken into account in this study:

1. Stability of ϵ^{VBF} with respect to τ_h-isolation: Maximum relative difference to weighted arithmetic mean at $\not{E}_\mathrm{T} \geq 30$ GeV. This amounts to $+17.26\%$ and -7.58% relative uncertainty.

2. Stability of ϵ^{VBF} with respect to $\not{E}_\mathrm{T}$-requirement: Relative difference to ϵ^{VBF} within uncertainties of weighted arithmetic mean at no selection on $\not{E}_\mathrm{T}$ as seen in Fig. 6.31. This amounts to $+8.30\%$ relative uncertainty on the upper edge and $+3.26\%$ on the lower edge.

Taking these new systematic uncertainties into account, including a shift of the central predicted value to accomodate for the bias due to $\not{E}_T$, the final background estimate with all systematic uncertainties for $N_{\text{SR}}^{\text{QCD}}$ is:

$$N_{\text{SR}}^{\text{QCD}} = 7.59 \pm 0.92 \text{ (stat)} \pm_{-0.72}^{+1.38} \text{ (syst)} \tag{6.26}$$

After the thorough estimation of all uncertainties of the background prediction, the expected yield of backgrounds is to be compared to the yield in data in the following section.

6.3.3 Limits

The following results are quoted from Ref. [74]. In Tab. 6.17, all background contributions with their statistical, systematic and combined uncertainties, as well as the final yield in data are shown.

sample	events	$\sigma_{\text{statistic}}$	$\sigma_{\text{systematic}}$	σ_{combined}
DY+jets	0.037	± 0.015	± 0.019	± 0.024
W+jets	0.530	± 0.040	± 0.265	± 0.268
VV	0.110	± 0.065	± 0.055	± 0.085
ttbar	0.110	± 0.012	± 0.055	± 0.056
single top	0.036	± 0.0066	± 0.018	± 0.019
QCD	7.59	± 0.92	$^{+1.38}_{-0.72}$	$^{+1.66}_{-1.17}$
Higgs	0.0005	± 0.00072	± 0.00025	± 0.00076
total	8.41	± 0.92	$^{+1.41}_{-0.77}$	$^{+1.68}_{-1.20}$
data	9			

TABLE 6.17: Final yields for all backgrounds and data for the LS $\tau_h \tau_h jj$ channel

The total expected background events and the yield of events in data are fully compatible within statistical and systematic uncertainties, as shown in Fig. 6.32.
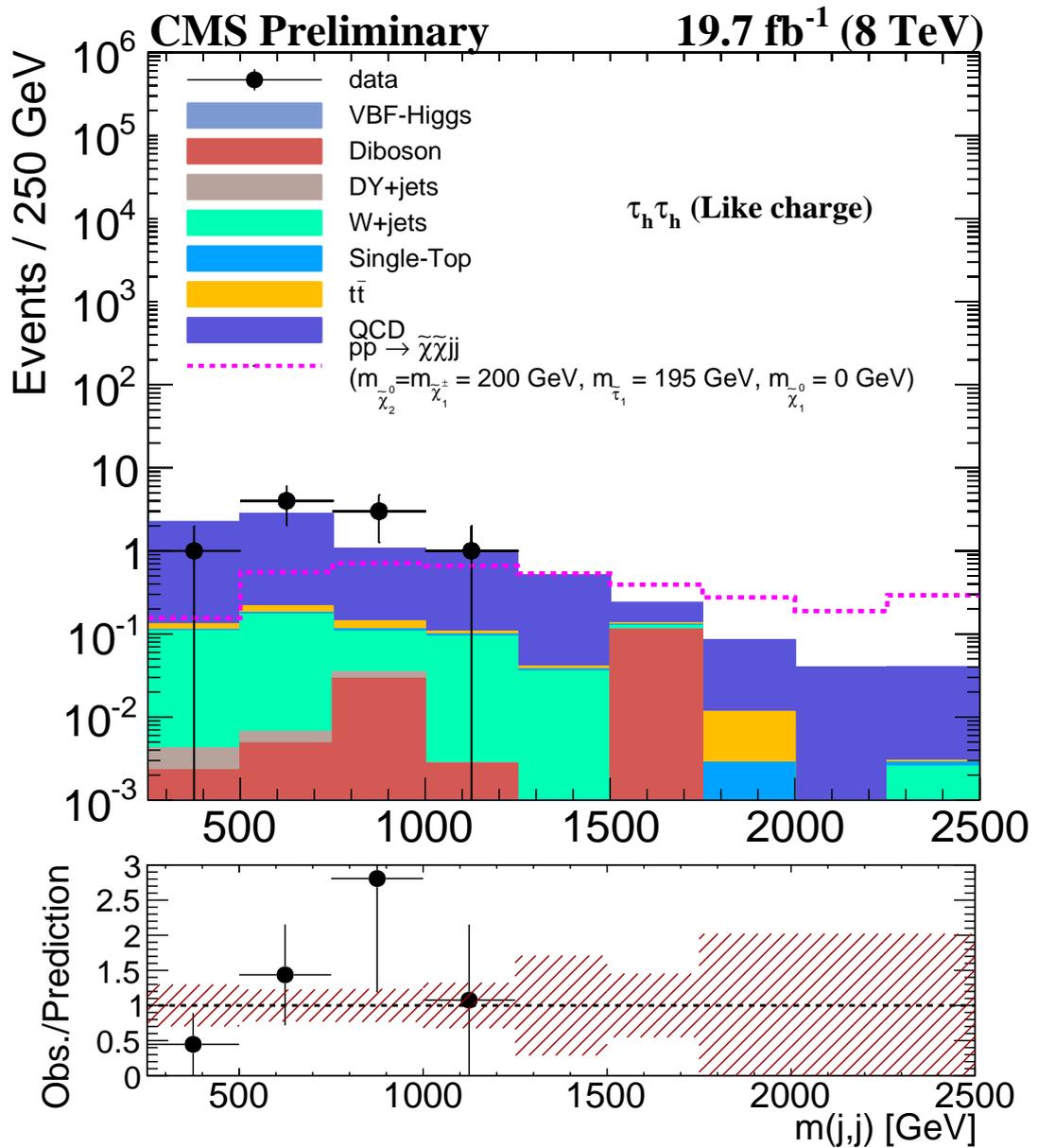
FIGURE 6.32: From Ref. [1]: The final yields of the like sign di-τ_h channel. No excess above standard model expectation is observed.

Therefore, limits on VBF SUSY have been set.

In Ref. [74], the limits are set on simplified models with a fixed LSP mass of $m_{\tilde\chi_1^0} = 0$ GeV and a variable chargino mass of 100 GeV $\leq m_{\tilde\chi_1^\pm} \leq$ 300 GeV with a $\tilde\tau$ of 95% of the chargino mass. Furthermore, the masses of the second neutralino $\tilde\chi_2^0$ and the chargino $\tilde\chi_1^\pm$ are set to the same value, as these gauginos belong to the same gauge group multiplet Ref. [1]. The interpretation considers the $\tilde\tau$ to be the lightest slepton while the colored sparticles are considered to be too massive to be produced. Therefore, the branching

ratios of the chargino and the second neutralino to stau are considered to be one. This interpretation results in the limits in Fig. 6.33.
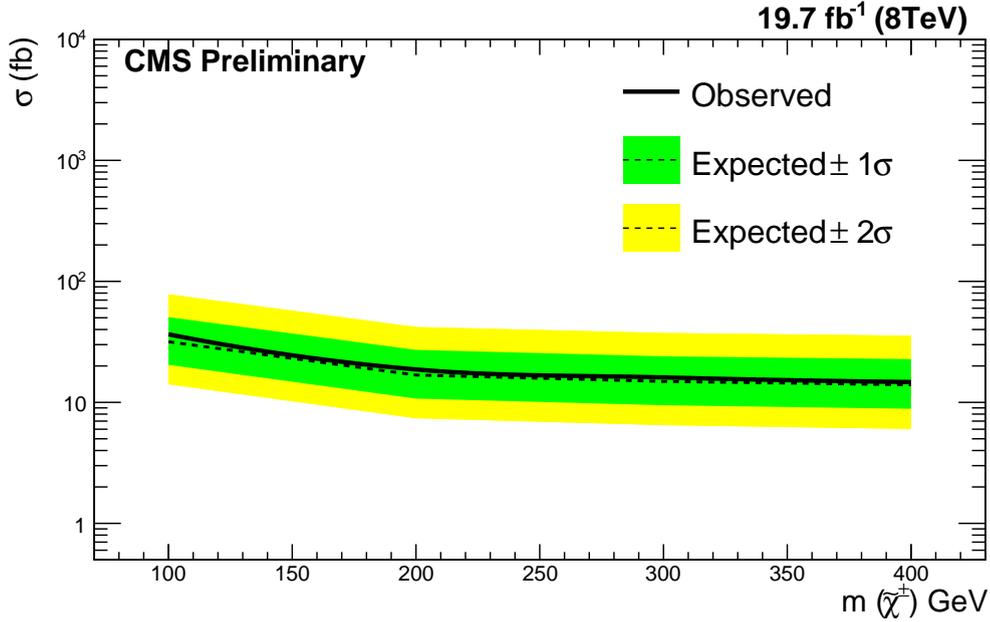
FIGURE 6.33: From Ref. [74]: Upper limit at the 95% CL on the cross-section as a function of $m_{\tilde{\chi}_2^0} = m_{\tilde{\chi}_1^{\pm}}$ with $m_{\tilde{\chi}_1^0} = 0$ for the LS $\tau_h\tau_h jj$ final state. The bands represents the one and two standard deviations obtained from the background-only hypothesis.

Comparing this limit specific to the LS di-τ_h channel to the overall limit in Fig. 2.10 that sets restraints on compressed spectrum SUSY up to 170 GeV and for negligible LSP mass up to 270 GeV shows how much the limitations set by the trigger also hampered this analysis. In essence, the LS di-τ_h channel cannot exclude any compressed spectrum SUSY and could only exclude chargino masses up to 150 GeV. The best-performing channel by far is the LS di-μ channel whose trigger has the highest efficiency and lowest momentum requirement.

6.3.4 Outline of data-driven improvements

The current analysis has three main impediments:

1. The trigger: It triggers on the τ_h and requires high thresholds. Thus, a trigger acceptance correction by isolation is needed and the limit set by the analysis is worsened due to the high τ_h momenta required.

2. The generator: Instead of using PYTHIA, MADGRAPH should be used for making QCD background predictions. As the lowest order process is already a four jet process, MADGRAPH might be a substantial improvement, especially regarding the VBF efficiency.

3. The fake probabilities: While the diced simulation yields a mostly self-consistent description, there is no guarantee that the flavor mixture of simulation corresponds to the same flavor mixture in data or that the isolation is simulated well. Indeed, there are hints to the contrary. The fact that OS and LS have different amounts of events is not simulated at all hints at some properties of the τ_h identification that are not simulated.

To amend the first problem, a jet-based trigger could be chosen for future analyses. The second problem can potentially be reduced by using MADGRAPH to generate more precise matrix elements for the four or more jets kinematics analysed. More challenging is the third kind of problem to get more accurate and reliable fake probabilities.

Using a dijet selection similar to the one in Sec. 5.5.2, it is possible to select very pure QCD events. It might be sufficient to select any reconstructed τ_h out of such a pure QCD selection, in order to derive probability maps already including the correct flavor mixture from data for dijets. Of course, this is only possible with sufficient statistics, limited to medium and high transverse momenta due to the trigger thresholds on jets. This introduces one further problem. Once those probability maps with the correct flavor combination for dijets are derived, one would need to correct this for the different flavor mixture of quadjet events that dominate the VBF analysis selection.

Even if such a correction would be hard to facilitate and might need a differential measurement of the fake probability in jet transverse momentum, in order to disentangle the different flavor contributions, it might be capable of reducing the systematic uncertainties due to τ_h isolation, substantially. Furthermore, the systematic uncertainty due to $\not{E}_T$ stability can be derived more reliably, if the fake probabilities were known from data itself.

Chapter 7

Conclusions and Outlook

In the CMS collaboration, factorized jet energy corrections (JEC) data are employed for data and simulation. The JEC for data include corrections from simulated events and data-driven techniques to derive residual differences between the jet energy response in data and simulation. In the barrel region of the detector ($|\eta| < 1.3$), the jet energy scale of these residual differences can be determined with high precision using $\gamma + $ jet or $Z^0 + $ jet events. This procedure makes use of the best-determined available reference objects in the form of leptons or photons to compare the jet energy against. The downside is that the phasespace covered by such methods is limited in transverse momentum and pseudorapidity, so it has to be extended. The best-performing candidate for such an extension is the use of dijet events that have significantly higher cross sections than the aforementioned processes. Although both objects have a nonnegligible energy resolution, facilitating such a dijet calibration as an extension of a $Z^0/\gamma + $ jet based method allows the coverage of the complete detector bar the most forward regions.

This thesis documents the relative residual (L2res) calibration with dijets for data taken with the CMS experiment at $\sqrt{s} = 8$ TeV in detail. The final results will also be published in Ref. [63].

Using the missing transverse momentum projection fraction method (MPF), corrections have been derived and refined during and after the 2012 datataking. In comparison to the $\sqrt{s} = 7$ TeV corrections, the granularity of the corrections has been increased down to calorimeter tower size in η due to the superior statistics available at $\sqrt{s} = 8$ TeV. Moreover, residual transverse momentum dependencies of the residual corrections as well as significant asymmetries with respect to $\text{sign}(\eta)$ have been observed and taken into account for the first time.

Despite a multifold increase of the granularity and complexity of the L2res correction, uncertainties for jets with $p_T \approx 100$ GeV have been reduced to sub percent level in $|\eta| < 2.5$. An exception are the endcaps and hadronic forward regions ($|\eta| \geq 2.5$) where

triggering on jet transverse momentum is less effective due to the momentum projection on the transverse plane scaling worse than in the central part of the detector. Meanwhile, the trigger prescales are constant in $|\eta|$ and have a major impact on the rate of forward jets being recorded that translate to low event counts for the jet energy corrections in the aforementioned forward regions.

For the next run at $\sqrt{s} = 13$ TeV, new forward jet triggers with $|\eta|$-dependent prescales are implemented that will allow a similar reduction in uncertainties for the forward regions as was observed with respect to the central region. Furthermore, the superior integrated luminosity expected will allow a ϕ-dependent correction in addition to the now implemented η- and $\bar{p}_T$-dependent correction. It is even possible to imagine a usage of the change in jet flavor composition as modeled in simulation depending on average dijet transverse momentum $\bar{p}_T$ to develop and implement flavor based jet energy corrections. But such an endeavor would necessitate very low uncertainties on the correction as well as an understanding of the origin of the observed $\bar{p}_T$-dependence.

Even today, jet energy corrections impact most analyses conducted at LHC experiments, given that they either use jets or use vetoes on jets. Further understanding and advances in this area will allow a large number of Standard Model (SM) and beyond SM (BSM) analyses to reduce their uncertainties and therefore increase their impact.

Although the SM is tried and tested to utmost precision in many areas, we already know that it has to be incomplete and Supersymmetry (SUSY) is a promising extension of the SM. In this thesis, a contribution to the first search for compressed spectrum SUSY in vector boson fusion (VBF) production channels (see Ref. [1]) is presented. For the like sign (LS) di-hadronic τ (τ_h) final state, a particularly challenging background composition has to be estimated. With 90% purity, multijet backgrounds being partly misidentified as τ_h constitute the main background. The rest of backgrounds is dominated by $W + 3$ jet events with one jet being misidentified as a τ_h. The dominant multijet background is estimated from data using an ABCD method utilizing inversion of VBF jet selection requirements and τ_h isolation requirements.

Studying systematic effects for the data-driven estimation of multijet background events on standard simulated events has not been possible. Out of $\mathcal{O} \approx 10^8$ events, the number of events left after all selections for multijet background is zero due to low simulated trigger rates and very low misidentification rates of jets as τ_h entering in quadrature. Determining expected properties of the τ_h response allows the redefinition of jet objects to τ_h objects within each reweighted event.

Using this "dicing" method allows to estimate otherwise not estimable systematic uncertainties to a level of 20% relative total uncertainty on the multijet background prediction. This uncertainty is fully competitive with less challenging final states like di-μ results.

Since no excess from the SM expectation has been observed, limits have been set. The VBF search was interpreted in a scenario where $\tilde{\chi}_1^\pm$ or $\tilde{\chi}_2^0$ mediated VBF $\tilde{\tau}$ production

ends in two to four τ being produced at a 100% branching ratio. Such scenarios with mass-degenerate $\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^0$, and $\tilde{\tau}$ are motivated by dark matter relic density models. The LS di-τ_h final state contributed to the final limit of $M(\tilde{\chi}_1^\pm = \tilde{\chi}_2^0) < 170$ GeV for a compressed scenario of $M(\tilde{\chi}_1^\pm) - M(\tilde{\chi}_1^0) = 50$ GeV at 95% confidence level (CL). Another interpretation for a massless LSP excludes $M(\tilde{\chi}_1^\pm = \tilde{\chi}_2^0) < 270$ GeV at 95% CL.

A trigger on τ_h had to be used that put too stringent requirements on τ_h transverse momentum. Therefore, despite the best branching ratio of the τ decay to hadronic decay products, the overall contribution of the LS di-τ_h final state to the combined exclusion limit is small. For the next run, utilizing forward jet triggers tagging VBF jets instead of soft leptons allows to probe compressed spectra even more thoroughly. As systematic uncertainties are already competititve for the LS di-τ_h final state, the superior branching ratio of τ decaying hadronically makes it imaginable that, given a trigger on VBF jets, this final state might be a significant contributor to a future VBF search for SUSY.

Appendix A

L2res correction control plots

A.1   Correction values

In this appendix chapter, all *L2res* corrections are shown with their explicit numeric values in the bins they are derived in and for the $\overline{p}_T$ range they are valid for. The formula used for the correction is stated in Eq. A.1.

$$L2res(p_T, \eta) = k_{\mathrm{FSR}}(|\eta|) \cdot [k_{\mathrm{const}}(\eta) + k_{p_T}(\eta) \cdot \log(\frac{\max(p_{T\,\min}, \min(p_{T\,\max}, p_T))}{1 \text{ GeV}})] \quad \text{(A.1)}$$

lower bin edge	upper bin edge	p_T min	p_T max	k_{FSR}	k_{const}	k_{p_T}
-5.191	-3.139	60	242	1.01188	1.07282	0
-3.139	-2.964	60	242	1.00677	1.13959	0
-2.964	-2.853	60	311	1.00380	1.20102	-0.0261331
-2.853	-2.5	60	468	0.995054	1.18864	-0.028539
-2.5	-2.322	60	628	0.996274	1.10072	-0.0116839
-2.322	-2.172	60	628	0.999032	1.0705	-0.00738168
-2.172	-2.043	60	628	0.998519	1.10048	-0.0142664
-2.043	-1.93	60	839	0.998427	1.10283	-0.0146567
-1.93	-1.83	60	839	1.00012	1.07255	-0.00916463
-1.83	-1.74	60	839	0.999506	1.06327	-0.00778776
-1.74	-1.653	60	839	0.998086	1.07915	-0.0119125
-1.653	-1.566	60	1121	0.998906	1.06561	-0.0104944
-1.566	-1.479	60	1121	0.998713	1.08072	-0.0134313
-1.479	-1.392	60	1121	0.998329	1.08267	-0.0145531
-1.392	-1.305	60	1121	0.999136	1.06249	-0.0109217
-1.305	-1.218	60	1121	1.00052	1.03226	-0.00452432
-1.218	-1.131	60	1121	0.999927	1.01069	-0.00113687
-1.131	-1.044	60	1121	1.00022	1.01735	-0.00239536
-1.044	-0.957	60	1497	1.00052	0.983895	0.00386298
-0.957	-0.879	60	1497	1.00045	0.981862	0.00423487
-0.879	-0.783	60	1497	1.00016	0.987009	0.00299773
-0.783	-0.696	60	1497	0.999511	1.0048	0.000714503
-0.696	-0.609	60	1497	0.999533	1.00849	-0.00120673
-0.609	-0.522	60	1497	0.999588	1.00327	-0.000797066
-0.522	-0.435	60	1497	0.999968	0.999752	-0.000444902
-0.435	-0.348	60	1497	1.00011	0.988306	0.00143032
-0.348	-0.261	60	1497	1.00037	0.983287	0.0020382
-0.261	-0.174	60	1497	1.00042	0.993188	0.0000457273
-0.174	-0.087	60	1497	1.00017	0.987095	0.000917325
-0.087	0	60	1497	1.0004	0.996132	-0.000840418

TABLE A.1: Correction values of the L2res correction for ak5PF jets for the negative pseudorapidity range of the detector.

lower bin edge	upper bin edge	$p_{T\,\text{min}}$	$p_{T\,\text{max}}$	k_{FSR}	k_{const}	k_{p_T}
0	0.087	60	1497	1.0004	0.985007	0.00135859
0.087	0.174	60	1497	1.00017	0.97973	0.00234734
0.174	0.261	60	1497	1.00042	0.99024	0.000573094
0.261	0.348	60	1497	1.00037	0.987179	0.00147179
0.348	0.435	60	1497	1.00011	1.00128	-0.000868817
0.435	0.522	60	1497	0.999968	1.00319	-0.000824581
0.522	0.609	60	1497	0.999588	1.00869	-0.0014742
0.609	0.696	60	1497	0.999533	1.00241	0.00028356
0.696	0.783	60	1497	0.999511	1.00735	-0.000330513
0.783	0.879	60	1497	1.00016	0.990618	0.00255055
0.879	0.957	60	1497	1.00045	0.984851	0.00386287
0.957	1.044	60	1497	1.00052	0.985942	0.00390898
1.044	1.131	60	1121	1.00022	1.00053	0.000924761
1.131	1.218	60	1121	0.999927	1.01578	-0.0015908
1.218	1.305	60	1121	1.00052	1.01419	-0.00189934
1.305	1.392	60	1121	0.999136	1.05728	-0.0101899
1.392	1.479	60	1121	0.998329	1.0819	-0.0142785
1.479	1.566	60	1121	0.998713	1.08871	-0.0155913
1.566	1.653	60	1121	0.998906	1.07741	-0.0137788
1.653	1.74	60	839	0.998086	1.06802	-0.0103594
1.74	1.83	60	839	0.999506	1.04936	-0.00643959
1.83	1.93	60	839	1.00012	1.05937	-0.00809463
1.93	2.043	60	839	0.998427	1.05733	-0.00752314
2.043	2.172	60	628	0.998519	1.06958	-0.00916676
2.172	2.322	60	628	0.999032	1.10479	-0.0135635
2.322	2.5	60	628	0.996274	1.11473	-0.0154937
2.5	2.853	60	468	0.995054	1.21048	-0.031416
2.853	2.964	60	311	1.00380	1.10887	0.000211617
2.964	3.139	60	242	1.00677	1.13959	0
3.139	5.191	60	242	1.01188	1.07282	0

TABLE A.2: Correction values of the L2res correction for ak5PF jets for the positive pseudorapidity range of the detector.

lower bin edge	upper bin edge	$p_{T\,\text{min}}$	$p_{T\,\text{max}}$	k_{FSR}	k_{const}	k_{p_T}
-5.191	-3.139	62	242	1.01011	1.06771	0
-3.139	-2.964	62	242	1.00246	1.13462	0
-2.964	-2.853	62	310	1.00227	1.12568	-0.0135338
-2.853	-2.5	62	467	0.996227	1.014932	-0.0222418
-2.5	-2.322	62	628	0.997426	1.08515	-0.00975238
-2.322	-2.172	62	628	0.999378	1.0756	-0.00945525
-2.172	-2.043	62	628	0.998359	1.09497	-0.0140674
-2.043	-1.93	62	839	0.998544	1.09445	-0.0138166
-1.93	-1.83	62	839	1.00012	1.06823	-0.0089031
-1.83	-1.74	62	839	0.999168	1.05019	-0.00613124
-1.74	-1.653	62	839	0.998496	1.07149	-0.0110065
-1.653	-1.566	62	1121	0.999376	1.057	-0.00930379
-1.566	-1.479	62	1121	0.998863	1.09085	-0.0152619
-1.479	-1.392	62	1121	0.998733	1.07411	-0.0132692
-1.392	-1.305	62	1121	0.999251	1.05854	-0.0103793
-1.305	-1.218	62	1121	1.00053	1.02409	-0.00324308
-1.218	-1.131	62	1121	0.999829	1.00412	-0.0000789537
-1.131	-1.044	62	1121	1.00032	1.01351	-0.00181314
-1.044	-0.957	62	1497	1.00067	0.988805	0.00308284
-0.957	-0.879	62	1497	1.00031	0.977472	0.00494603
-0.879	-0.783	62	1497	1.00021	0.991388	0.00232894
-0.783	-0.696	62	1497	0.999551	1.00287	0.000345489
-0.696	-0.609	62	1497	0.999627	1.00879	-0.00116421
-0.609	-0.522	62	1497	0.999594	0.999298	-0.000175439
-0.522	-0.435	62	1497	1.00009	0.99664	0.000052436
-0.435	-0.348	62	1497	1.00035	0.983588	0.00214248
-0.348	-0.261	62	1497	1.00029	0.989091	0.00112916
-0.261	-0.174	62	1497	1.00039	0.9973746	-0.000782938
-0.174	-0.087	62	1497	0.999903	0.984925	0.00124099
-0.087	0	62	1497	1.00026	0.997768	-0.00109646

TABLE A.3: Correction values of the L2res correction for ak5PFchs jets for the negative pseudorapidity range of the detector.

lower bin edge	upper bin edge	$p_{T\,\text{min}}$	$p_{T\,\text{max}}$	k_{FSR}	k_{const}	k_{p_T}
0	0.087	62	1497	1.00026	0.985357	0.00125972
0.087	0.174	62	1497	0.999903	0.979914	0.00232729
0.174	0.261	62	1497	1.00039	0.993175	0.000125716
0.261	0.348	62	1497	1.00029	0.989999	0.000994603
0.348	0.435	62	1497	1.00035	1.00137	-0.000865584
0.435	0.522	62	1497	1.00009	1.00296	-0.000803275
0.522	0.609	62	1497	0.999594	1.01641	-0.00271861
0.609	0.696	62	1497	0.999627	1.00643	-0.000371202
0.696	0.783	62	1497	0.999551	1.00659	-0.000150217
0.783	0.879	62	1497	1.00021	0.986046	0.00330973
0.879	0.957	62	1497	1.00031	0.976181	0.00534711
0.957	1.044	62	1497	1.00067	0.990991	0.00312511
1.044	1.131	62	1121	1.00032	1.00266	0.000468847
1.131	1.218	62	1121	0.999829	1.01338	-0.00111408
1.218	1.305	62	1121	1.00053	1.01651	-0.00227739
1.305	1.392	62	1121	0.999251	1.05957	-0.010621
1.392	1.479	62	1121	0.9998733	1.07707	-0.0136643
1.479	1.566	62	1121	0.998863	1.08157	-0.0145453
1.566	1.653	62	1121	0.999376	1.07417	-0.0135148
1.653	1.74	62	839	0.998496	1.06316	-0.00993686
1.74	1.83	62	839	0.999168	1.05542	-0.00783172
1.83	1.93	62	839	1.00012	1.06684	-0.00988894
1.93	2.043	62	839	0.998544	1.05	-0.00693879
2.043	2.172	62	628	0.998359	1.06829	-0.00960675
2.172	2.322	62	628	0.999378	1.09169	-0.0123021
2.322	2.5	62	628	0.997426	1.1084	-0.0152169
2.5	2.853	62	467	0.996227	1.16975	-0.0245419
2.853	2.964	62	310	1.00227	1.13732	-0.00666699
2.964	3.139	62	242	1.00246	1.13462	0
3.139	5.191	62	242	1.01011	1.06771	0

TABLE A.4: Correction values of the L2res correction for ak5PFchs jets for the positive pseudorapidity range of the detector.

lower bin edge	upper bin edge	$p_{T\,\text{min}}$	$p_{T\,\text{max}}$	k_{FSR}	k_{const}	k_{p_T}
-5.191	-3.139	71	261	1.01404	1.07065	0
-3.139	-2.964	71	261	1.00099	1.13340	0
-2.964	-2.853	71	332	1.00578	1.17098	-0.0233715
-2.853	-2.5	71	494	0.995783	1.1819	-0.0280072
-2.5	-2.322	71	494	0.999261	1.10251	-0.0127784
-2.322	-2.172	71	628	0.999299	1.07655	-0.00967229
-2.172	-2.043	71	839	0.999493	1.06511	-0.00892468
-2.043	-1.93	71	839	0.999669	1.065811	-0.00893163
-1.93	-1.83	71	839	0.999696	1.04088	-0.0042452
-1.83	-1.74	71	839	0.999442	1.0375	-0.00405545
-1.74	-1.653	71	839	0.998693	1.06068	-0.00902998
-1.653	-1.566	71	1121	0.998774	10.6657	-0.0107757
-1.566	-1.479	71	1121	0.998926	1.09272	-0.0154478
-1.479	-1.392	71	1121	0.997898	1.07744	-0.0137382
-1.392	-1.305	71	1121	0.998913	1.06393	-0.0109262
-1.305	-1.218	71	1121	0.999708	1.02932	-0.00411256
-1.218	-1.131	71	1121	1.00005	1.01548	-0.00189026
-1.131	-1.044	71	1121	1.00056	1.01024	-0.00119427
-1.044	-0.957	71	1497	1.00021	0.984502	0.0037402
-0.957	-0.879	71	1497	0.99983	0.983728	0.00375376
-0.879	-0.783	71	1497	1.00032	0.992261	0.00214104
-0.783	-0.696	71	1497	0.99995	1.00507	-0.000123023
-0.696	-0.609	71	1497	0.999676	0.997169	0.000593513
-0.609	-0.522	71	1497	0.999518	1.00447	-0.00101459
-0.522	-0.435	71	1497	1.00014	0.994476	0.000421312
-0.435	-0.348	71	1497	0.999919	0.98965	0.00113511
-0.348	-0.261	71	1497	1.00056	0.992756	0.000594253
-0.261	-0.174	71	1497	1.00032	0.989479	0.000702306
-0.174	-0.087	71	1497	1.00035	0.985118	0.00129462
-0.087	0	71	1497	1.00025	1.00083	-0.00152927

TABLE A.5: Correction values of the L2res correction for ak7PF jets for the negative pseudorapidity range of the detector.

lower bin edge	upper bin edge	$p_{T\,\text{min}}$	$p_{T\,\text{max}}$	k_{FSR}	k_{const}	k_{p_T}
0	0.087	71	1497	1.00025	0.976007	0.00285721
0.087	0.174	71	1497	1.00035	0.978378	0.00251891
0.174	0.261	71	1497	1.00032	0.995993	-0.000236869
0.261	0.348	71	1497	1.00056	0.990844	0.000863273
0.348	0.435	71	1497	0.999919	0.9995	-0.000516811
0.435	0.522	71	1497	1.00014	0.995598	0.000377545
0.522	0.609	71	1497	0.999518	1.01497	-0.00254147
0.609	0.696	71	1497	0.999676	0.997957	0.000869225
0.696	0.783	71	1497	0.999995	0.995514	0.00153298
0.783	0.879	71	1497	1.00032	0.994676	0.00189323
0.879	0.957	71	1497	0.999983	0.9839	0.00399749
0.957	1.044	71	1497	1.00021	0.986412	0.0038513
1.044	1.131	71	1121	1.00056	1.01584	-0.00146728
1.131	1.218	71	1121	1.00005	1.02105	-0.00237568
1.218	1.305	71	1121	0.999708	1.03251	-0.00488892
1.305	1.392	71	1121	0.998913	1.06347	-0.0110435
1.392	1.479	71	1121	0.997898	1.0878	-0.0152466
1.479	1.566	71	1121	0.998926	1.09218	-0.0160734
1.566	1.653	71	1121	0.998774	1.07332	-0.0131336
1.653	1.74	71	839	0.998693	1.07012	-0.0108915
1.74	1.83	71	839	0.999442	1.05239	-0.00721019
1.83	1.93	71	839	0.999696	1.04708	-0.00630058
1.93	2.043	71	839	0.999669	1.04347	-0.00572395
2.043	2.172	71	839	0.999493	1.0457	-0.00566599
2.172	2.322	71	628	0.999299	1.06505	-0.00782059
2.322	2.5	71	494	0.999261	1.09032	-0.0119319
2.5	2.853	71	494	0.995783	1.21464	-0.033505
2.853	2.964	71	332	1.00578	0.807277	0.0562704
2.964	3.139	71	261	1.00099	1.13340	0
3.139	5.191	71	261	1.01404	1.07065	0

TABLE A.6: Correction values of the L2res correction for ak7PF jets for the positive pseudorapidity range of the detector.

A.2 Control plots - kFSR extrapolations

This section covers all kFSR extrapolations over α in all bins of $|\eta|$ used to derive the relative residual corrections for ak5PF, ak5PFCHS and ak7PF jets.

A.2.1 kFSR extrapolations - ak5PF

All extrapolations over α for ak5PF jets are shown. The correction is the ratio of the value at $\alpha < 0.2$ to the extrapolated value at $\alpha = 0$. Statistical correlations between bins are not taken into account, despite their inclusivity.
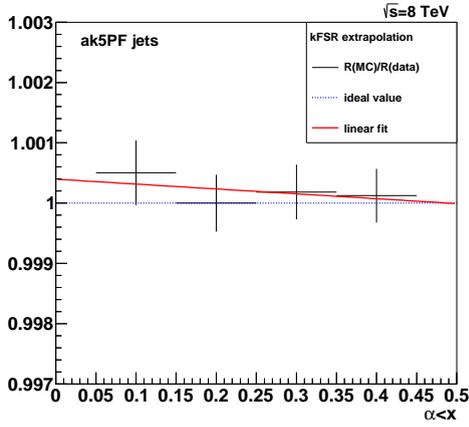
(a) $0.000 \leq |\eta| < 0.087$
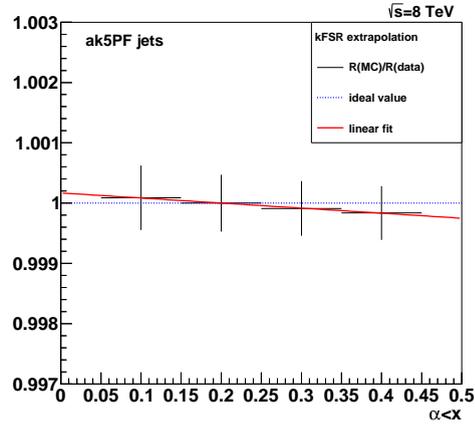
(b) $0.087 \leq |\eta| < 0.174$
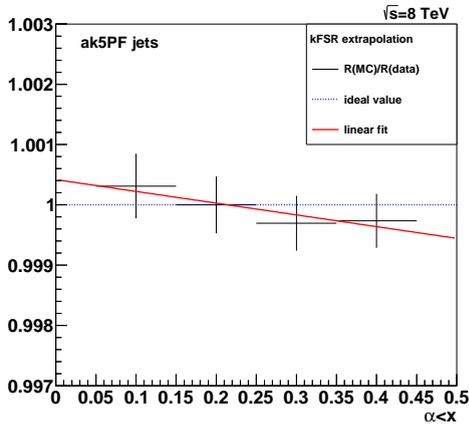
(c) $0.174 \leq |\eta| < 0.261$
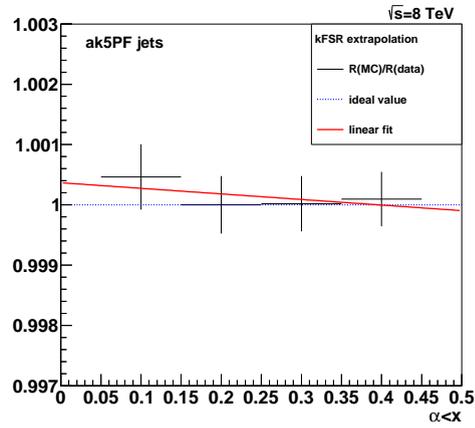
(d) $0.261 \leq |\eta| < 0.348$
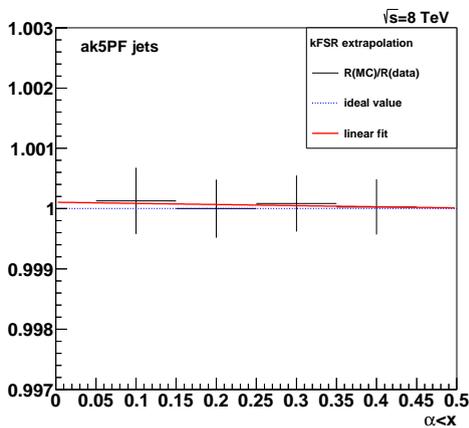
(e) $0.348 \leq |\eta| < 0.435$
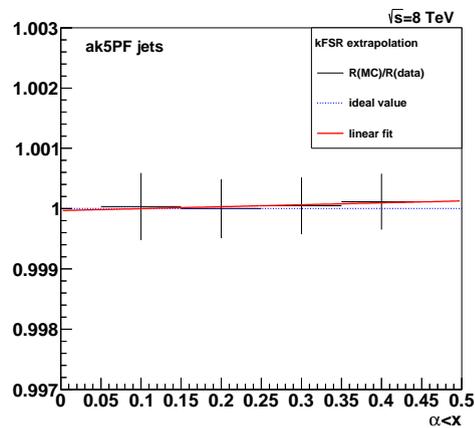
(f) $0.435 \leq |\eta| < 0.522$

FIGURE A.1: Linear kFSR extrapolations per $|\eta|$ bin over α in $0 \leq |\eta| < 0.522$ for ak5PF jets
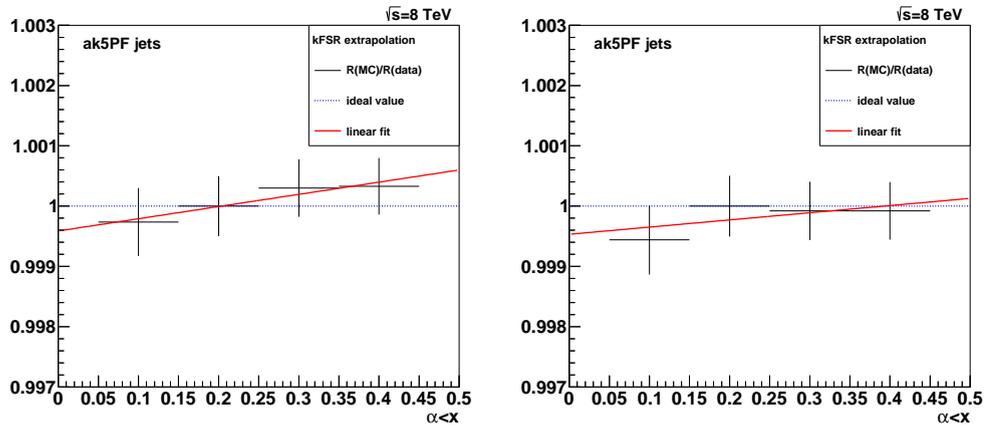
(a) $0.522 \leq |\eta| < 0.609$

(b) $0.609 \leq |\eta| < 0.696$
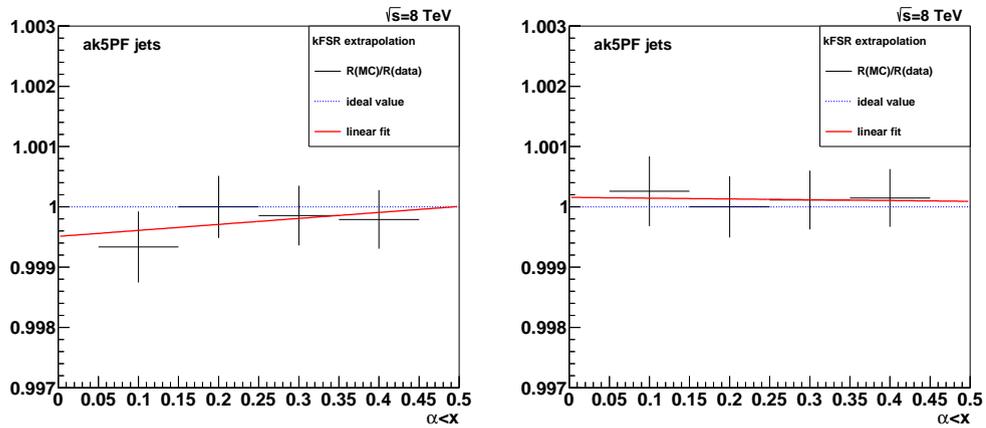
(c) $0.696 \leq |\eta| < 0.783$

(d) $0.783 \leq |\eta| < 0.879$
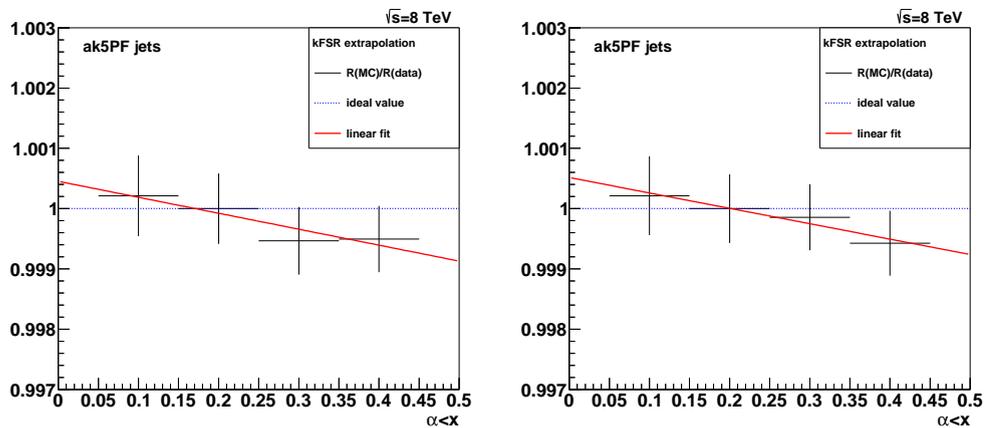
(e) $0.879 \leq |\eta| < 0.957$

(f) $0.957 \leq |\eta| < 1.044$

FIGURE A.2: Linear kFSR extrapolations per $|\eta|$ bin over α in $0.522 \leq |\eta| < 1.044$ for ak5PF jets
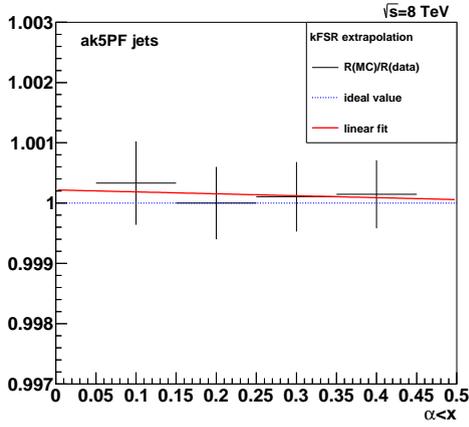
(a) $1.044 \leq |\eta| < 1.131$
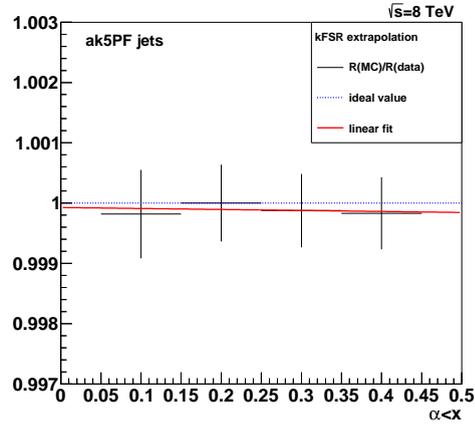
(b) $1.131 \leq |\eta| < 1.218$
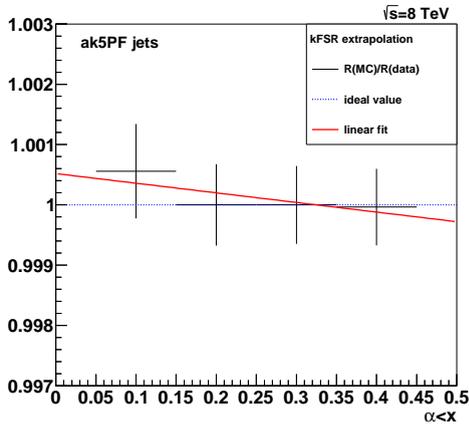
(c) $1.218 \leq |\eta| < 1.305$
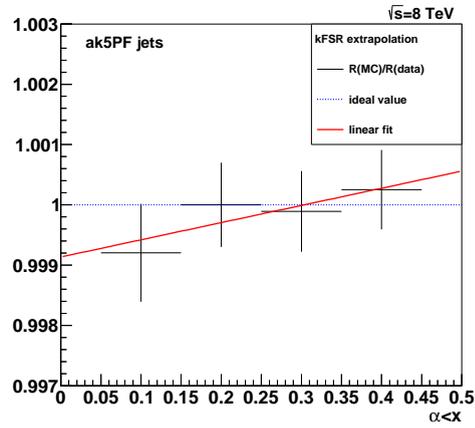
(d) $1.305 \leq |\eta| < 1.392$
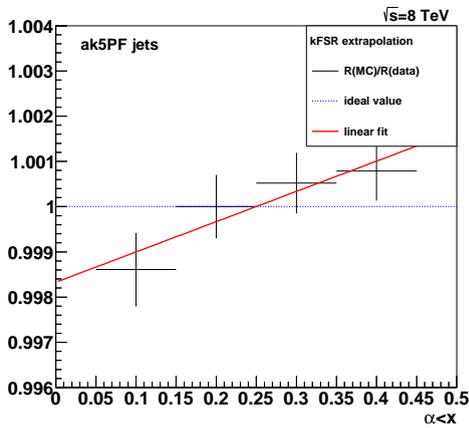
(e) $1.392 \leq |\eta| < 1.479$
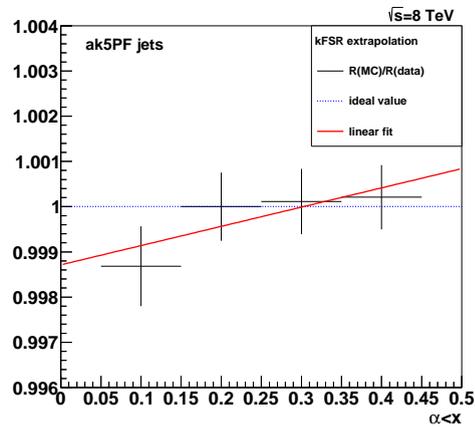
(f) $1.479 \leq |\eta| < 1.566$

FIGURE A.3: Linear kFSR extrapolations per $|\eta|$ bin over α in $1.044 \leq |\eta| < 1.566$ for ak5PF jets
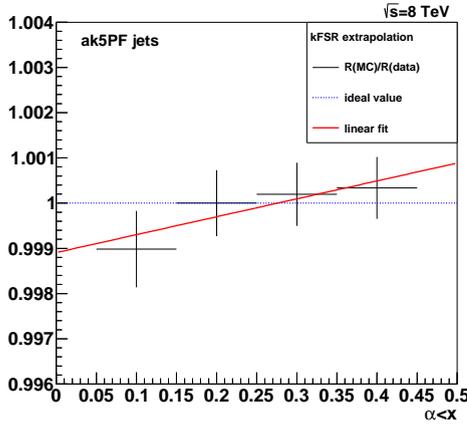
(a) $1.566 \leq |\eta| < 1.653$
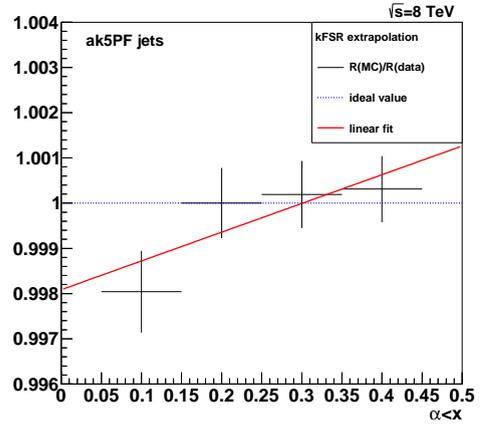
(b) $1.653 \leq |\eta| < 1.640$
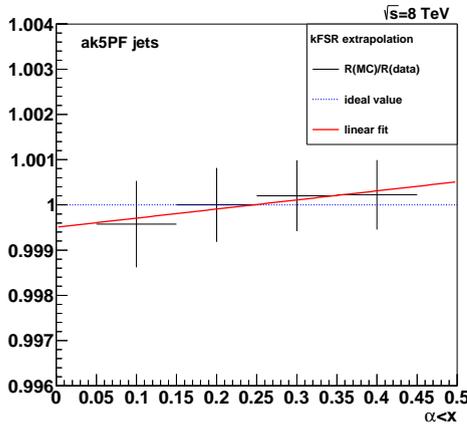
(c) $1.640 \leq |\eta| < 1.830$
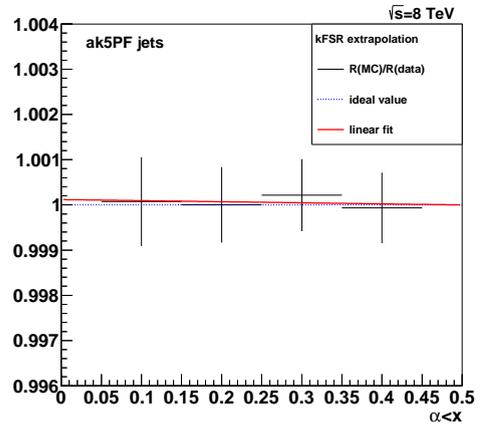
(d) $1.830 \leq |\eta| < 1.930$
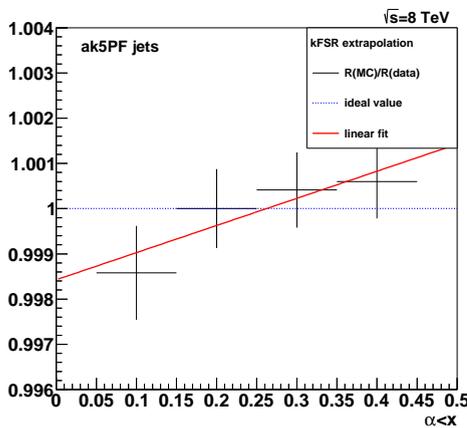
(e) $1.930 \leq |\eta| < 2.043$
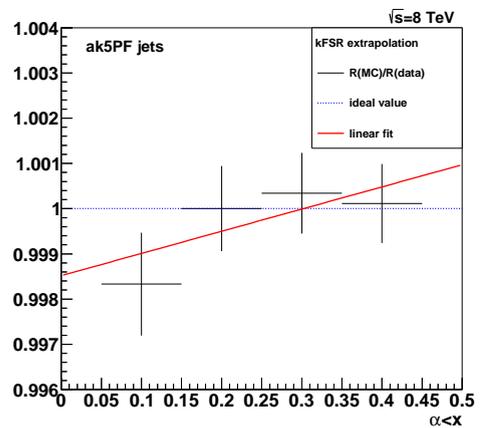
(f) $2.043 \leq |\eta| < 2.172$

FIGURE A.4: Linear kFSR extrapolations per $|\eta|$ bin over α in $1.566 \leq |\eta| < 2.172$ for ak5PF jets
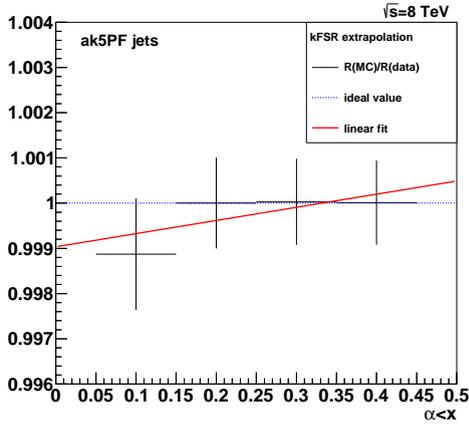
(a) $2.172 \leq |\eta| < 2.322$
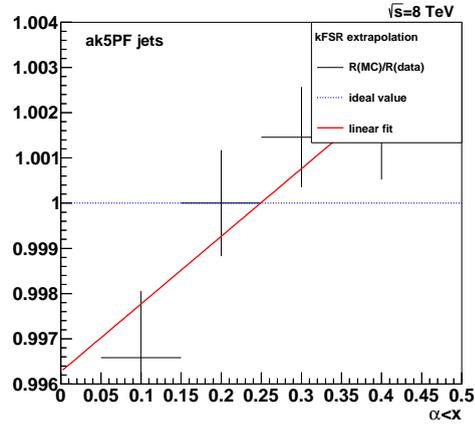
(b) $2.322 \leq |\eta| < 2.500$
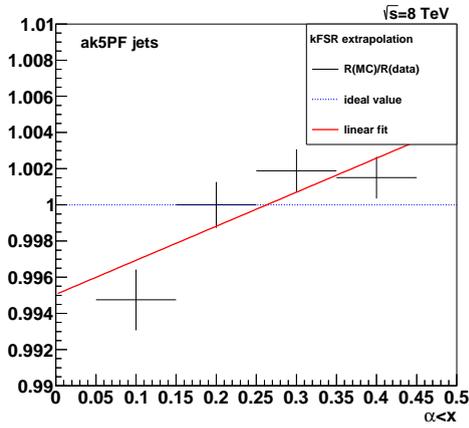
(c) $2.500 \leq |\eta| < 2.853$
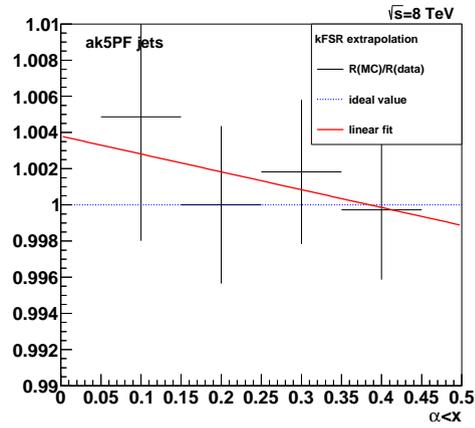
(d) $2.853 \leq |\eta| < 2.964$
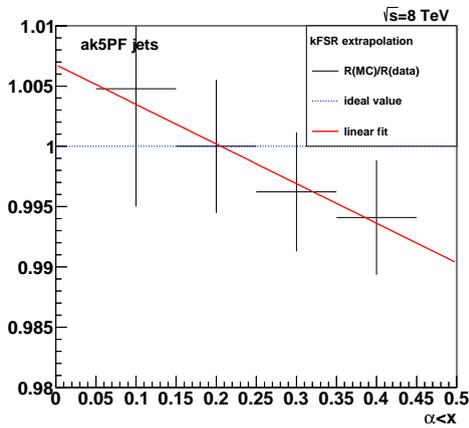
(e) $2.964 \leq |\eta| < 3.139$
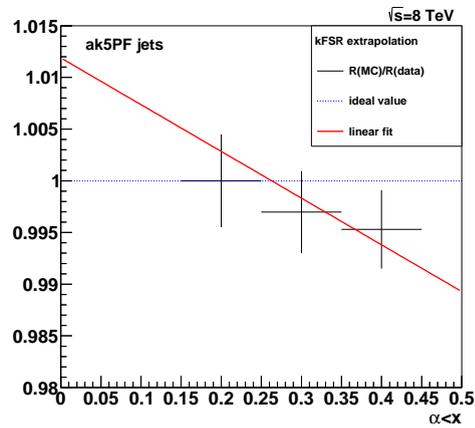
(f) $3.139 \leq |\eta| < 5.232$

FIGURE A.5: Linear kFSR extrapolations per $|\eta|$ bin over α in $2.172 \leq |\eta| < 5.232$ for ak5PF jets

A.2.2   kFSR extrapolations - ak5PF+CHS

All extrapolations over α for ak5PF+CHS jets are shown. The correction is the ratio of the value at $\alpha < 0.2$ to the extrapolated value at $\alpha = 0$. Statistical correlations between bins are not taken into account, despite their inclusivity.
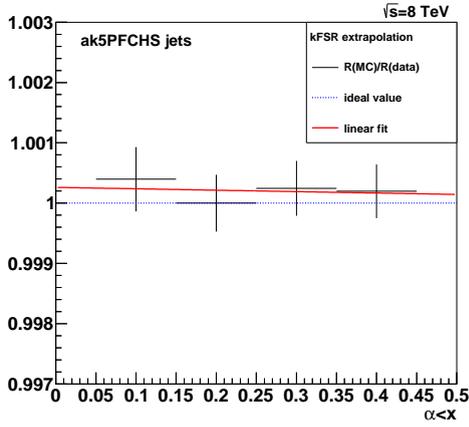
(a) $0.000 \leq |\eta| < 0.087$
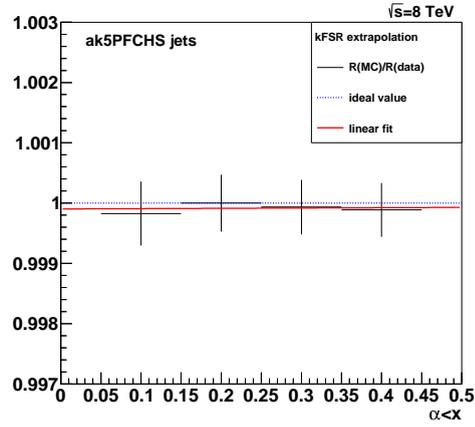
(b) $0.087 \leq |\eta| < 0.174$
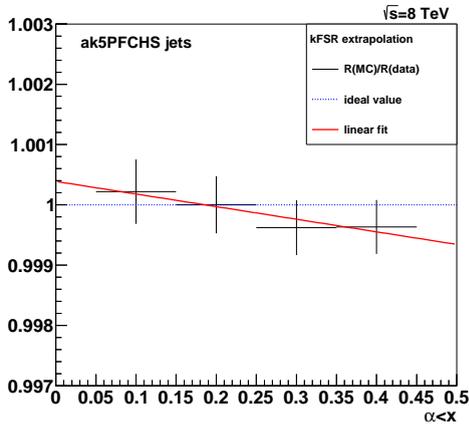
(c) $0.174 \leq |\eta| < 0.261$
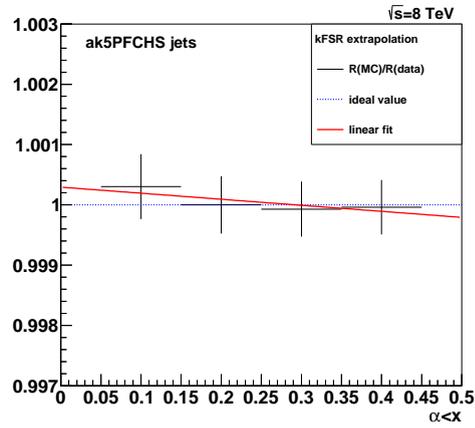
(d) $0.261 \leq |\eta| < 0.348$
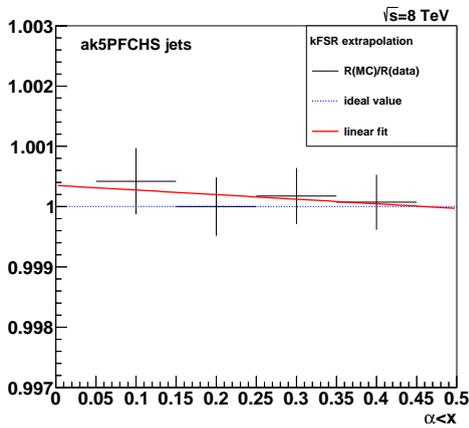
(e) $0.348 \leq |\eta| < 0.435$
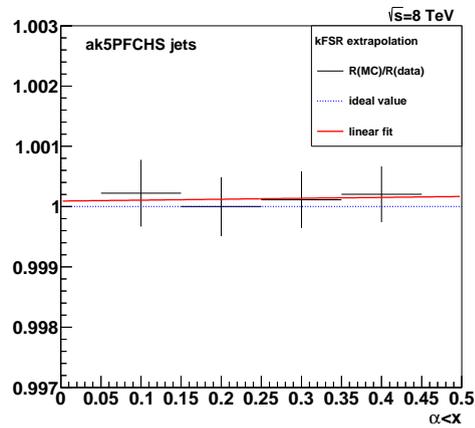
(f) $0.435 \leq |\eta| < 0.522$

FIGURE A.6: Linear kFSR extrapolations per $|\eta|$ bin over α in $0 \leq |\eta| < 0.522$ for ak5PF+CHS jets
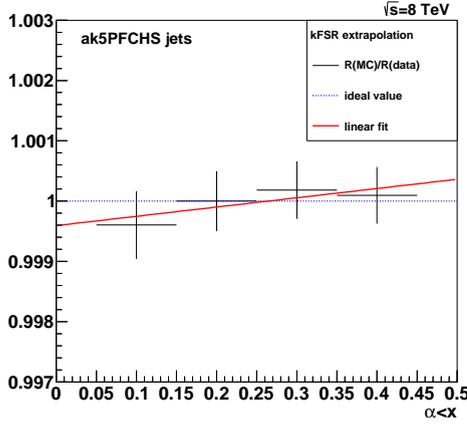
(a) $0.522 \leq |\eta| < 0.609$
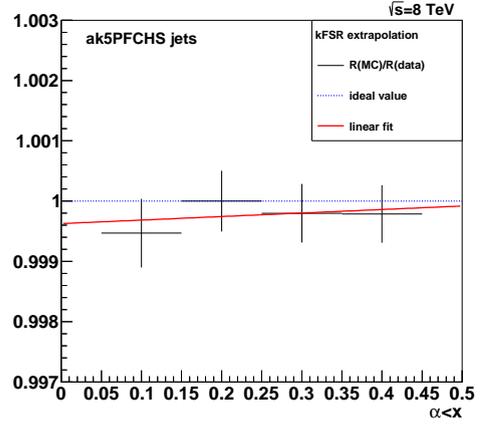
(b) $0.609 \leq |\eta| < 0.696$
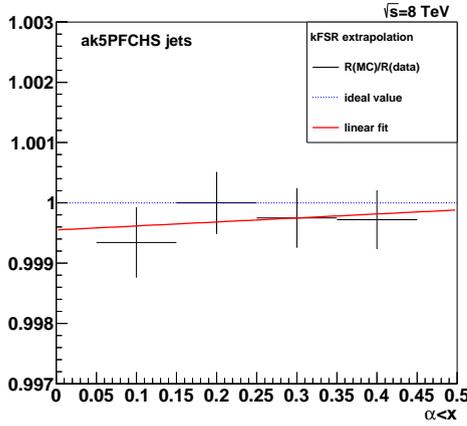
(c) $0.696 \leq |\eta| < 0.783$
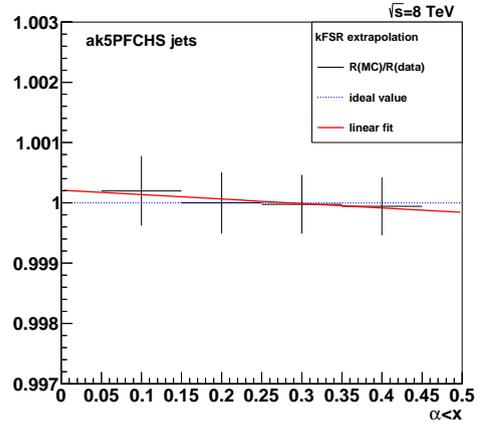
(d) $0.783 \leq |\eta| < 0.879$
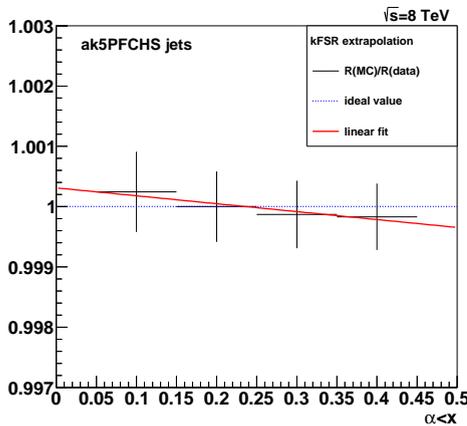
(e) $0.879 \leq |\eta| < 0.957$
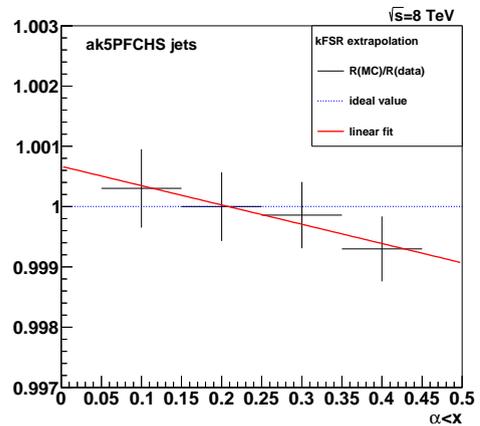
(f) $0.957 \leq |\eta| < 1.044$

FIGURE A.7: Linear kFSR extrapolations per $|\eta|$ bin over α in $0.522 \leq |\eta| < 1.044$ for ak5PF+CHS jets
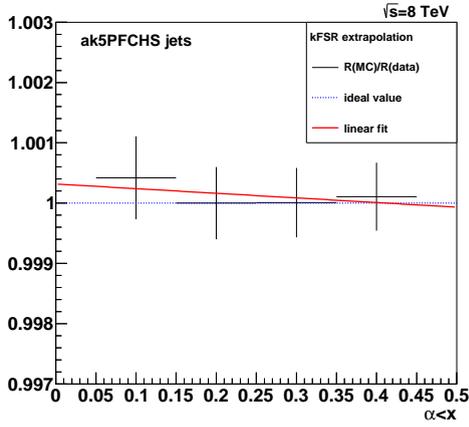
(a) $1.044 \leq |\eta| < 1.131$
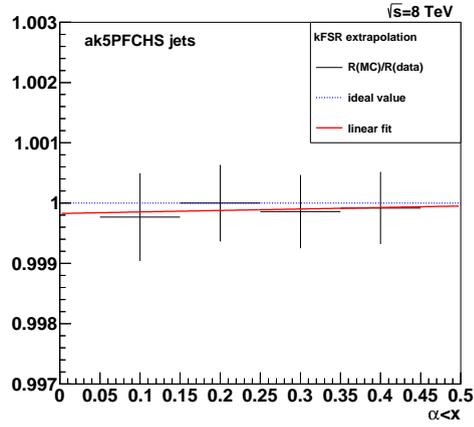
(b) $1.131 \leq |\eta| < 1.218$
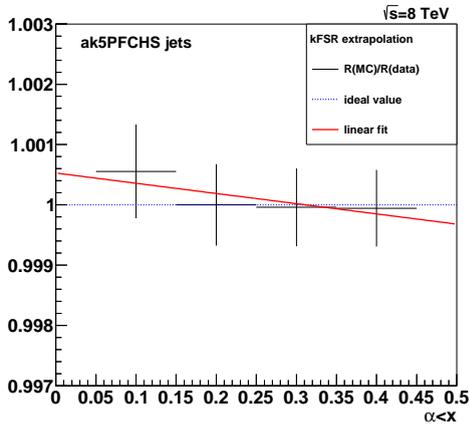
(c) $1.218 \leq |\eta| < 1.305$
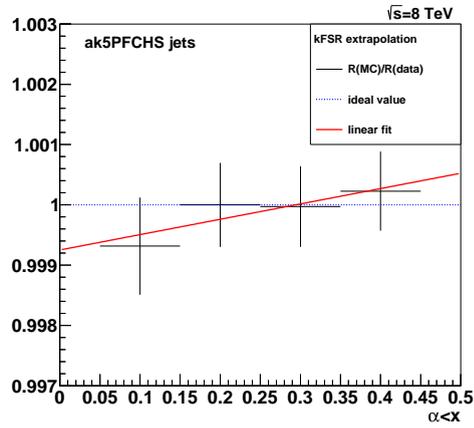
(d) $1.305 \leq |\eta| < 1.392$
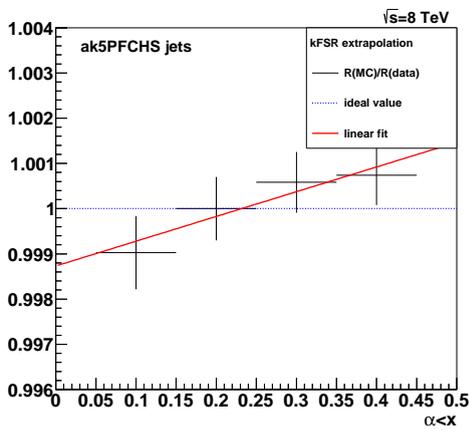
(e) $1.392 \leq |\eta| < 1.479$
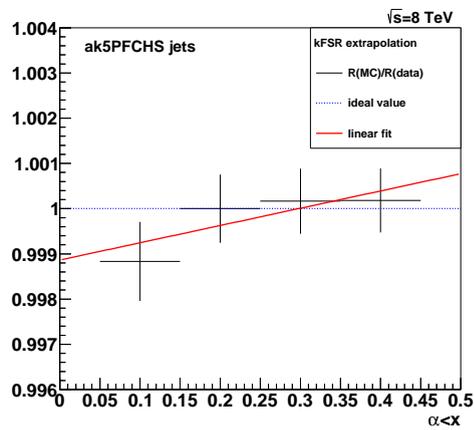
(f) $1.479 \leq |\eta| < 1.566$

FIGURE A.8: Linear kFSR extrapolations per $|\eta|$ bin over α in $1.044 \leq |\eta| < 1.566$ for ak5PF+CHS jets
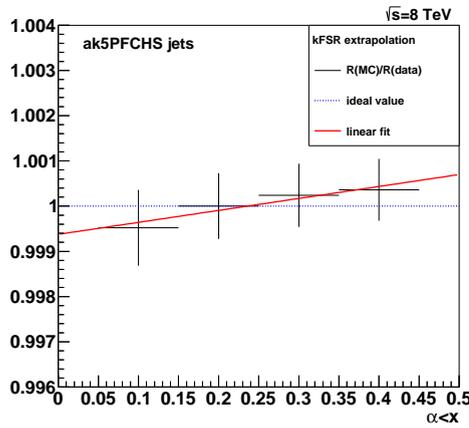
(a) $1.566 \leq |\eta| < 1.653$
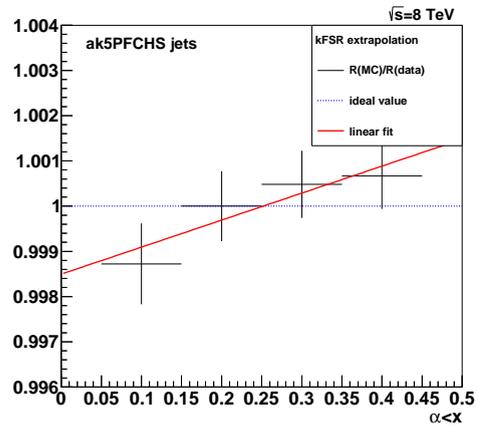
(b) $1.653 \leq |\eta| < 1.640$
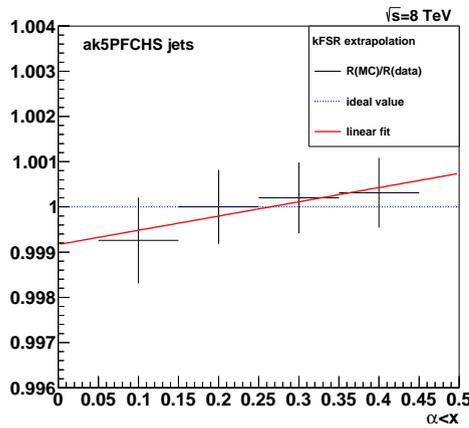
(c) $1.640 \leq |\eta| < 1.830$
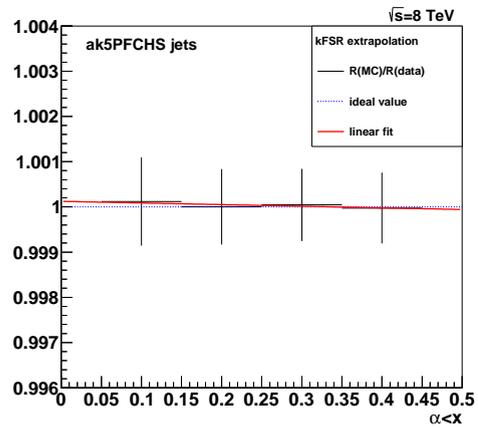
(d) $1.830 \leq |\eta| < 1.930$
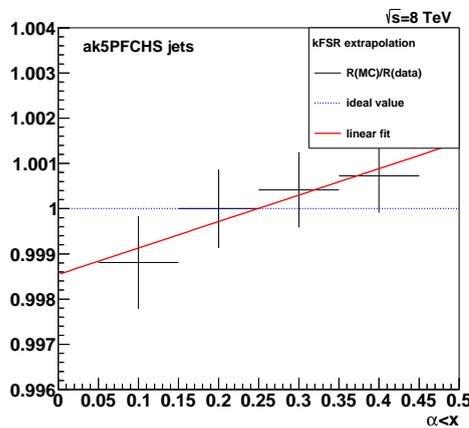
(e) $1.930 \leq |\eta| < 2.043$
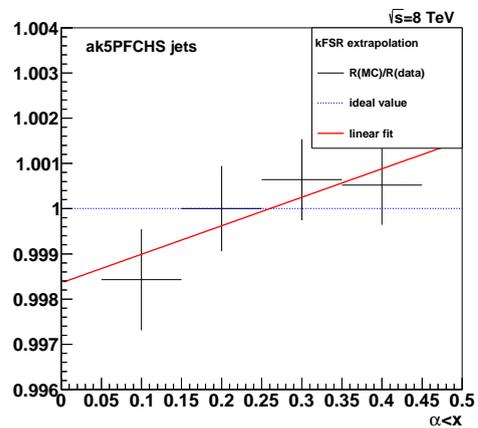
(f) $2.043 \leq |\eta| < 2.172$

FIGURE A.9: Linear kFSR extrapolations per $|\eta|$ bin over α in $1.566 \leq |\eta| < 2.172$ for ak5PF+CHS jets
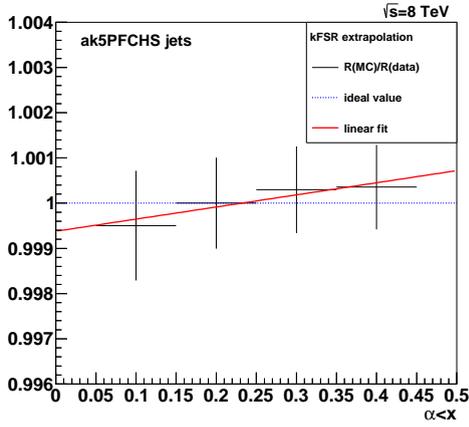
(a) $2.172 \leq |\eta| < 2.322$
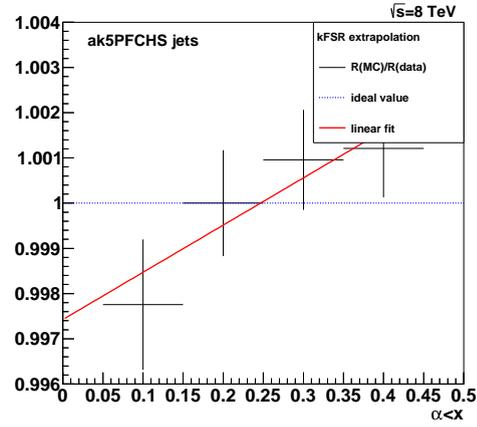
(b) $2.322 \leq |\eta| < 2.500$
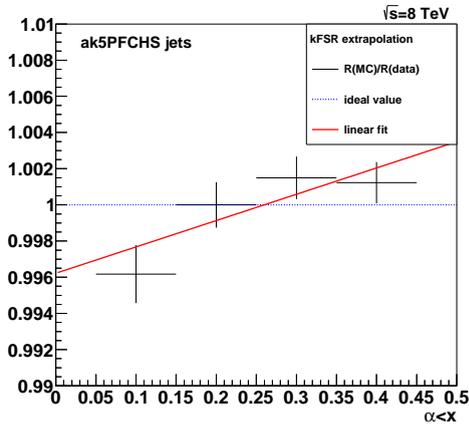
(c) $2.500 \leq |\eta| < 2.853$
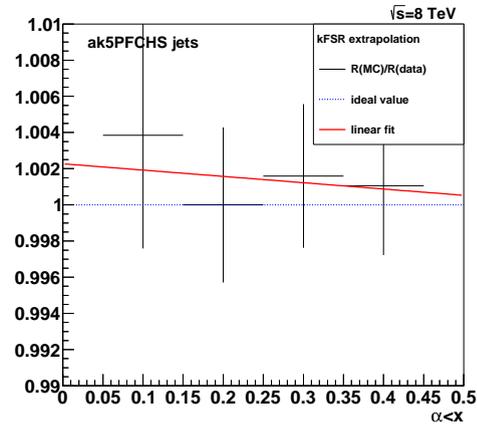
(d) $2.853 \leq |\eta| < 2.964$
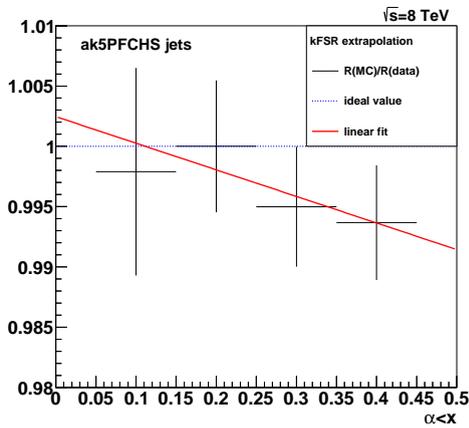
(e) $2.964 \leq |\eta| < 3.139$
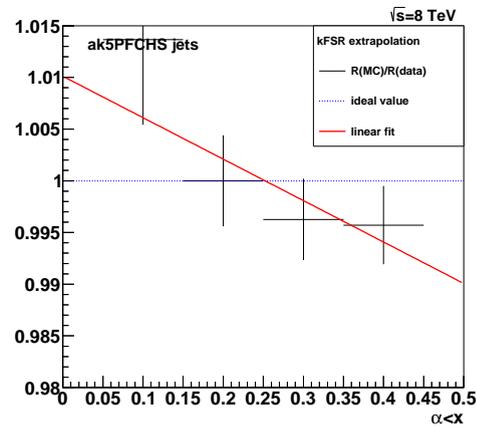
(f) $3.139 \leq |\eta| < 5.232$

FIGURE A.10: Linear kFSR extrapolations per $|\eta|$ bin over α in $2.172 \leq |\eta| < 5.232$ for ak5PF+CHS jets

A.2.3 kFSR extrapolations - ak7PF

All extrapolations over α for ak7PF jets are shown. The correction is the ratio of the value at $\alpha < 0.2$ to the extrapolated value at $\alpha = 0$. Statistical correlations between bins are not taken into account, despite their inclusivity.
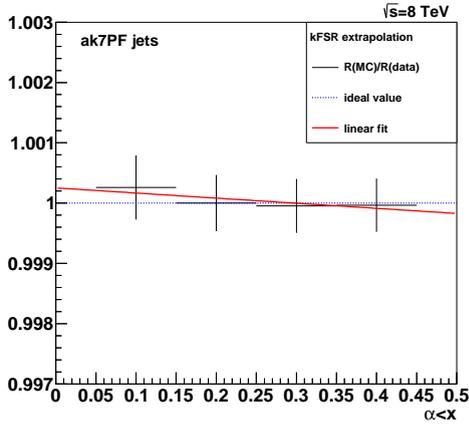
(a) $0.000 \leq |\eta| < 0.087$
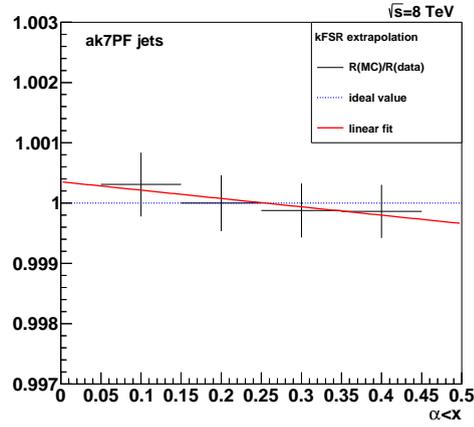
(b) $0.087 \leq |\eta| < 0.174$
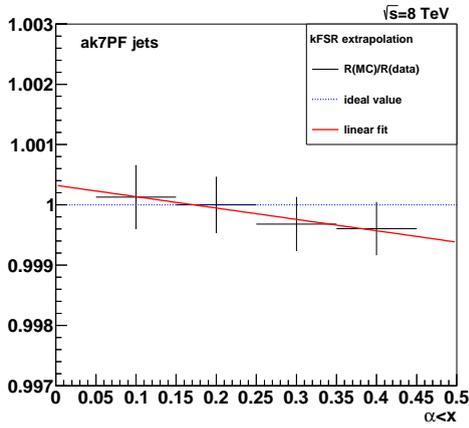
(c) $0.174 \leq |\eta| < 0.261$
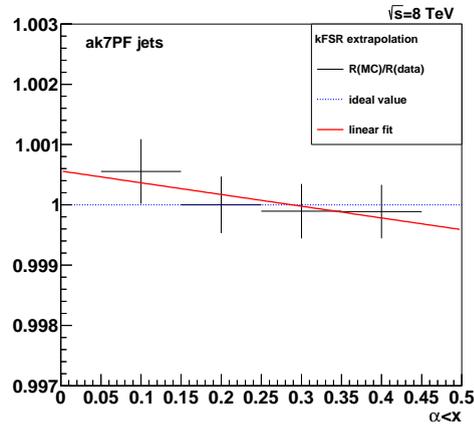
(d) $0.261 \leq |\eta| < 0.348$
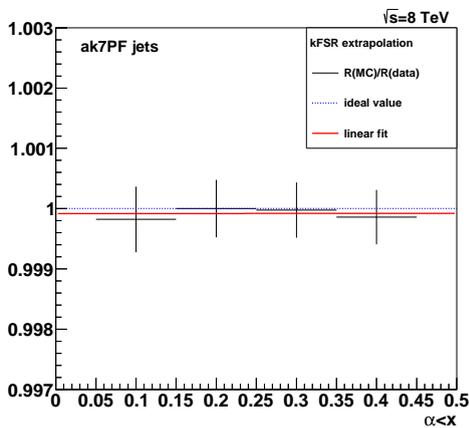
(e) $0.348 \leq |\eta| < 0.435$
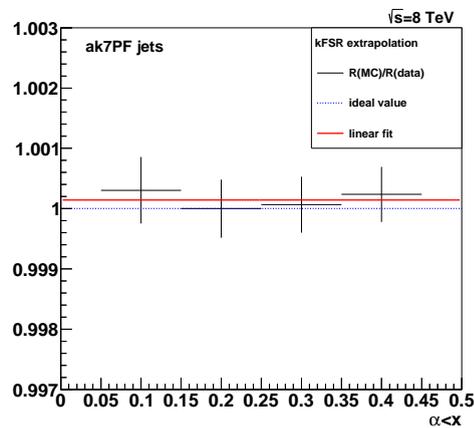
(f) $0.435 \leq |\eta| < 0.522$

FIGURE A.11: Linear kFSR extrapolations per $|\eta|$ bin over α in $0 \leq |\eta| < 0.522$ for ak7PF jets
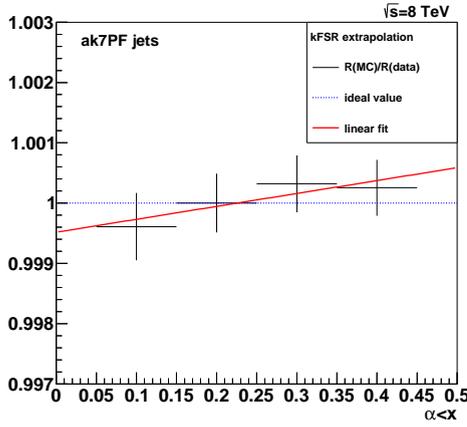
(a) $0.522 \leq |\eta| < 0.609$
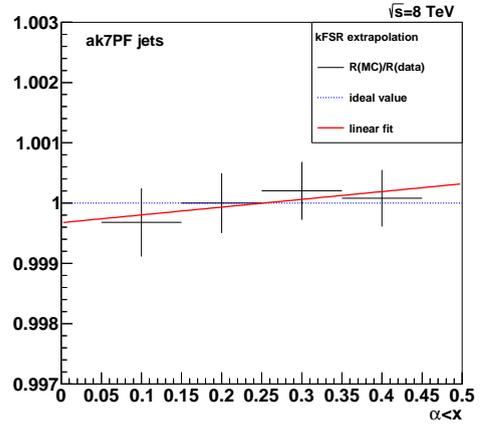
(b) $0.609 \leq |\eta| < 0.696$
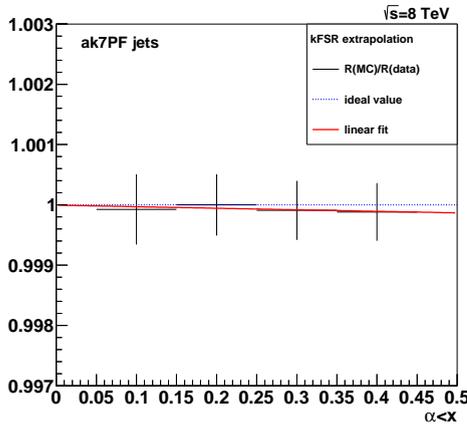
(c) $0.696 \leq |\eta| < 0.783$
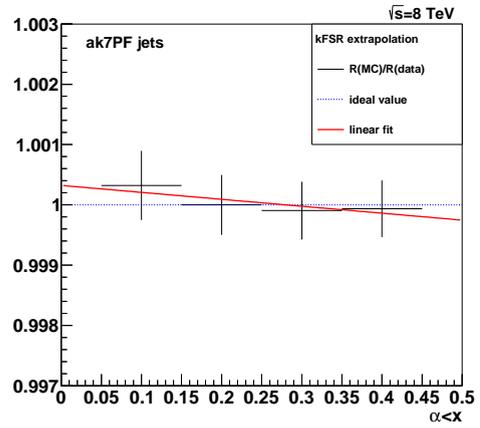
(d) $0.783 \leq |\eta| < 0.879$
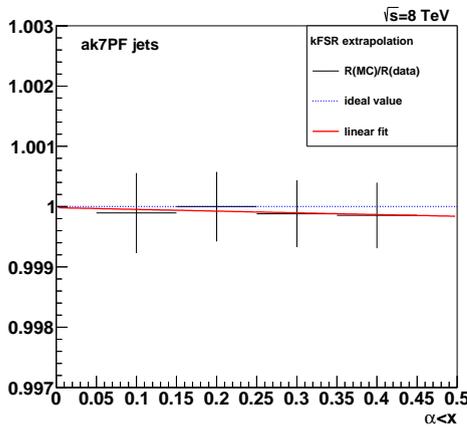
(e) $0.879 \leq |\eta| < 0.957$
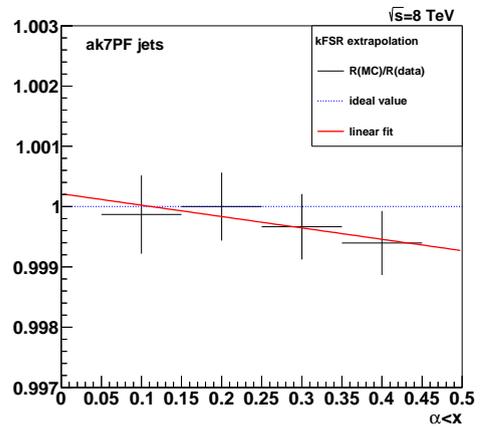
(f) $0.957 \leq |\eta| < 1.044$

FIGURE A.12: Linear kFSR extrapolations per $|\eta|$ bin over α in $0.522 \leq |\eta| < 1.044$ for ak7PF jets
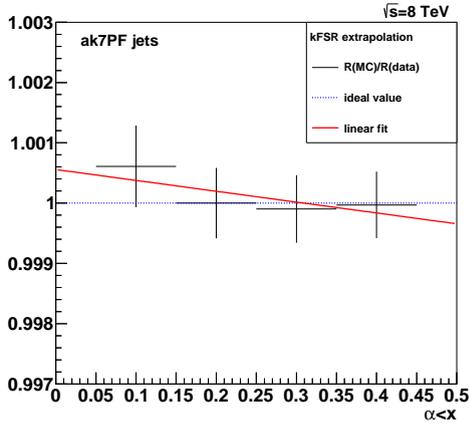
(a) $1.044 \leq |\eta| < 1.131$
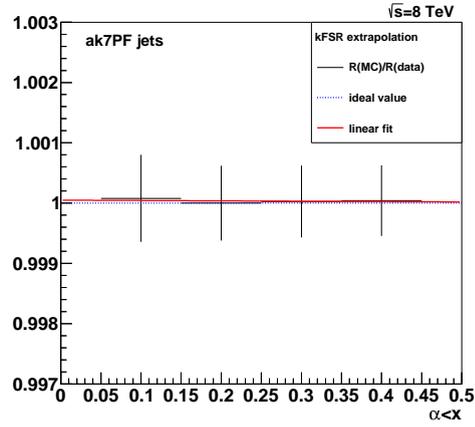
(b) $1.131 \leq |\eta| < 1.218$
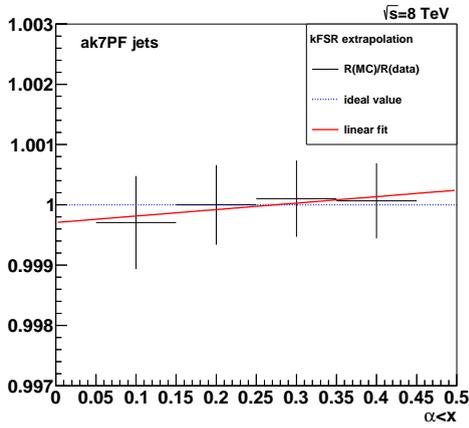
(c) $1.218 \leq |\eta| < 1.305$
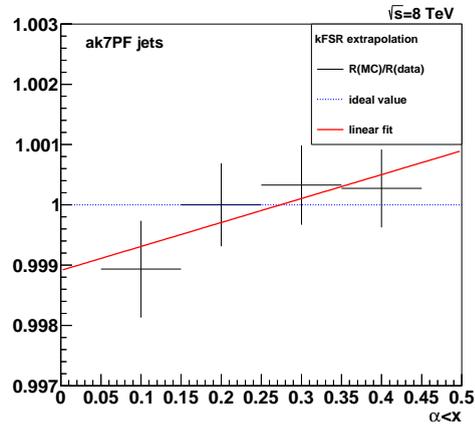
(d) $1.305 \leq |\eta| < 1.392$
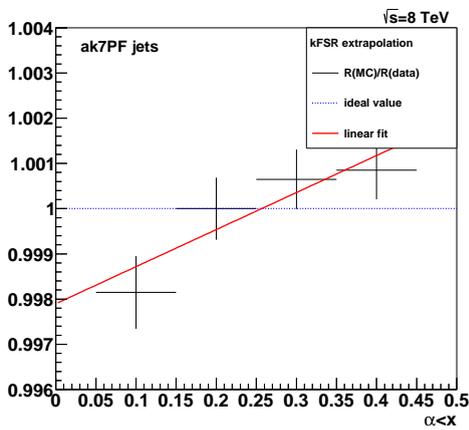
(e) $1.392 \leq |\eta| < 1.479$
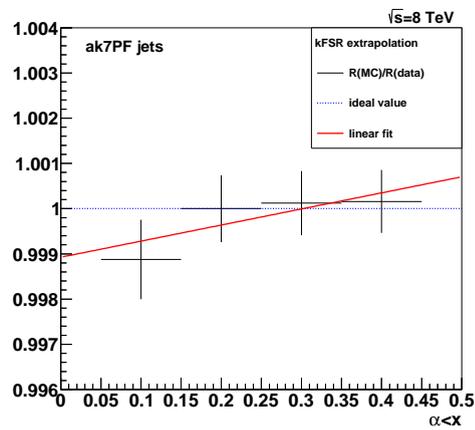
(f) $1.479 \leq |\eta| < 1.566$

FIGURE A.13: Linear kFSR extrapolations per $|\eta|$ bin over α in $1.044 \leq |\eta| < 1.566$ for ak7PF jets
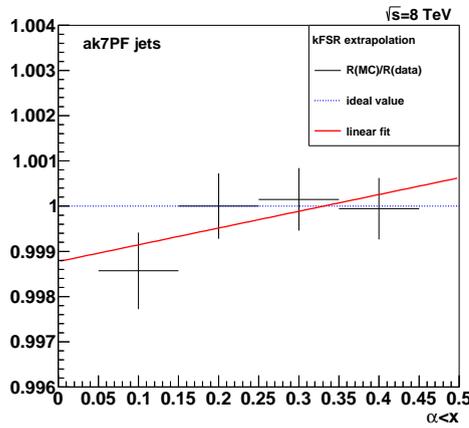
(a) $1.566 \leq |\eta| < 1.653$
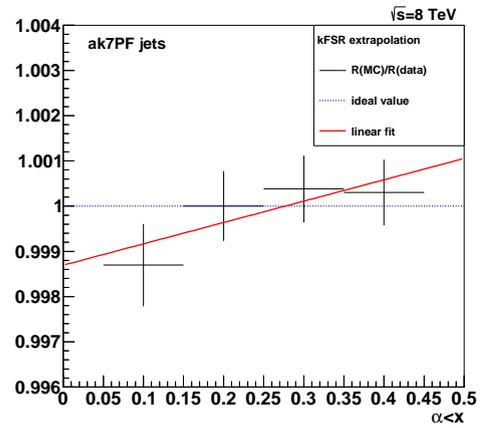
(b) $1.653 \leq |\eta| < 1.640$
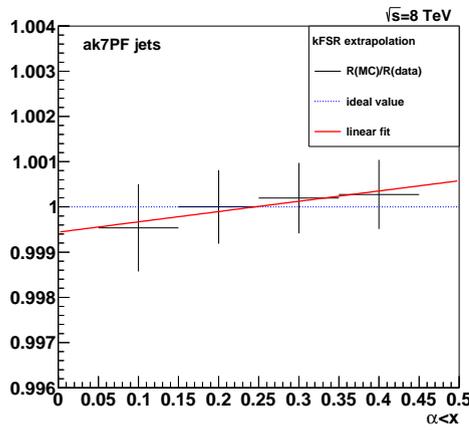
(c) $1.640 \leq |\eta| < 1.830$
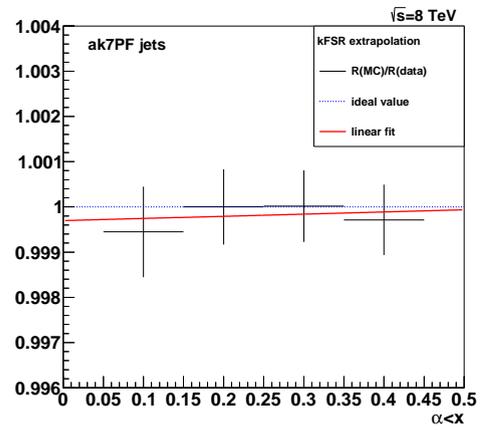
(d) $1.830 \leq |\eta| < 1.930$
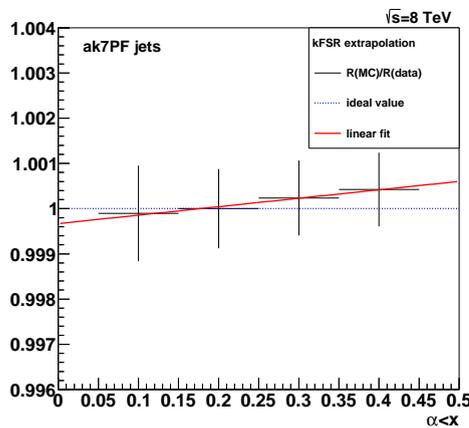
(e) $1.930 \leq |\eta| < 2.043$
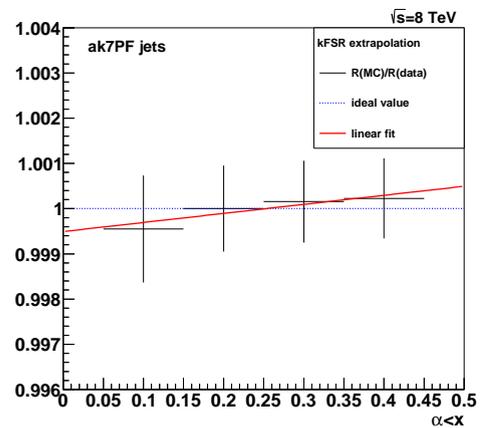
(f) $2.043 \leq |\eta| < 2.172$

FIGURE A.14: Linear kFSR extrapolations per $|\eta|$ bin over α in $1.566 \leq |\eta| < 2.172$ for ak7PF jets
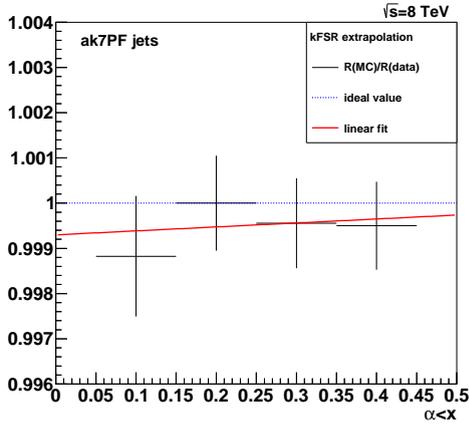
(a) $2.172 \leq |\eta| < 2.322$
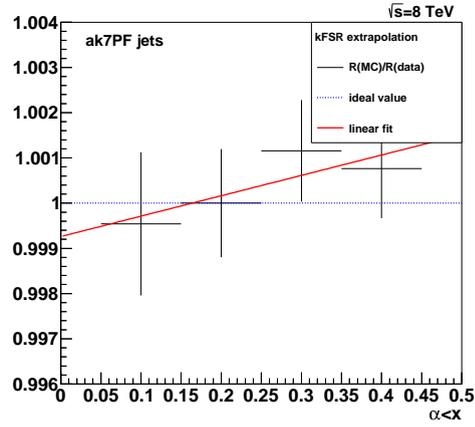
(b) $2.322 \leq |\eta| < 2.500$
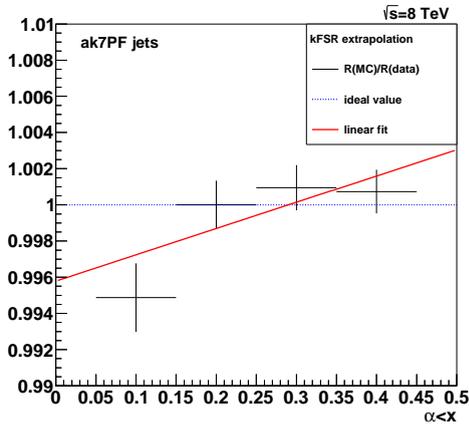
(c) $2.500 \leq |\eta| < 2.853$
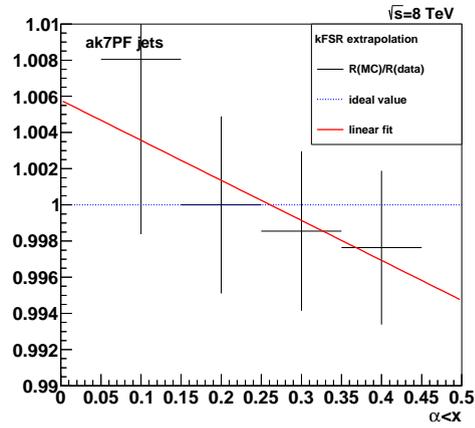
(d) $2.853 \leq |\eta| < 2.964$
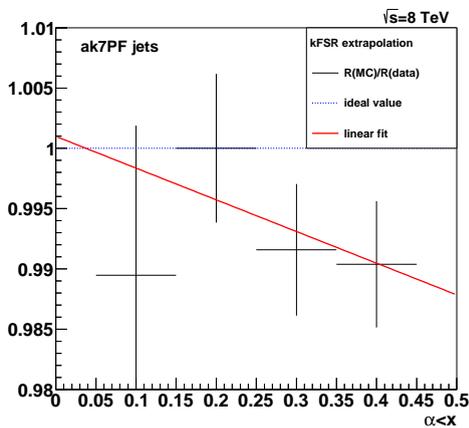
(e) $2.964 \leq |\eta| < 3.139$
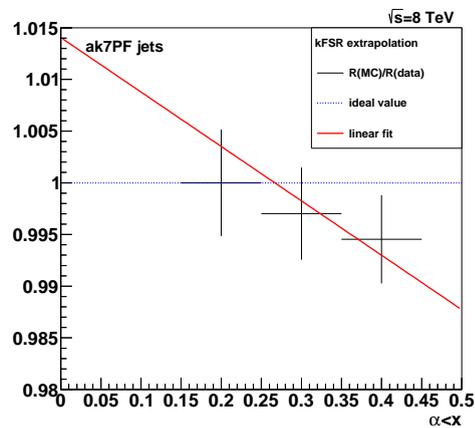
(f) $3.139 \leq |\eta| < 5.232$

FIGURE A.15: Linear kFSR extrapolations per $|\eta|$ bin over α in $2.172 \leq |\eta| < 5.232$ for ak7PF jets

A.3   Control plots - $\bar{p}_T$ extrapolations

This section covers all MPF extrapolations over $\bar{p}_T$ in all bins of η used to derive the relative residual corrections for ak5PF, ak5PFCHS and ak7PF jets.

A.3.1 $\bar{p}_T$ extrapolations - ak5PF

All extrapolations over $\bar{p}_T$ for ak5PF jets are shown in a constant and a loglinear ($[0] + [1] \cdot \log(x)$) version.
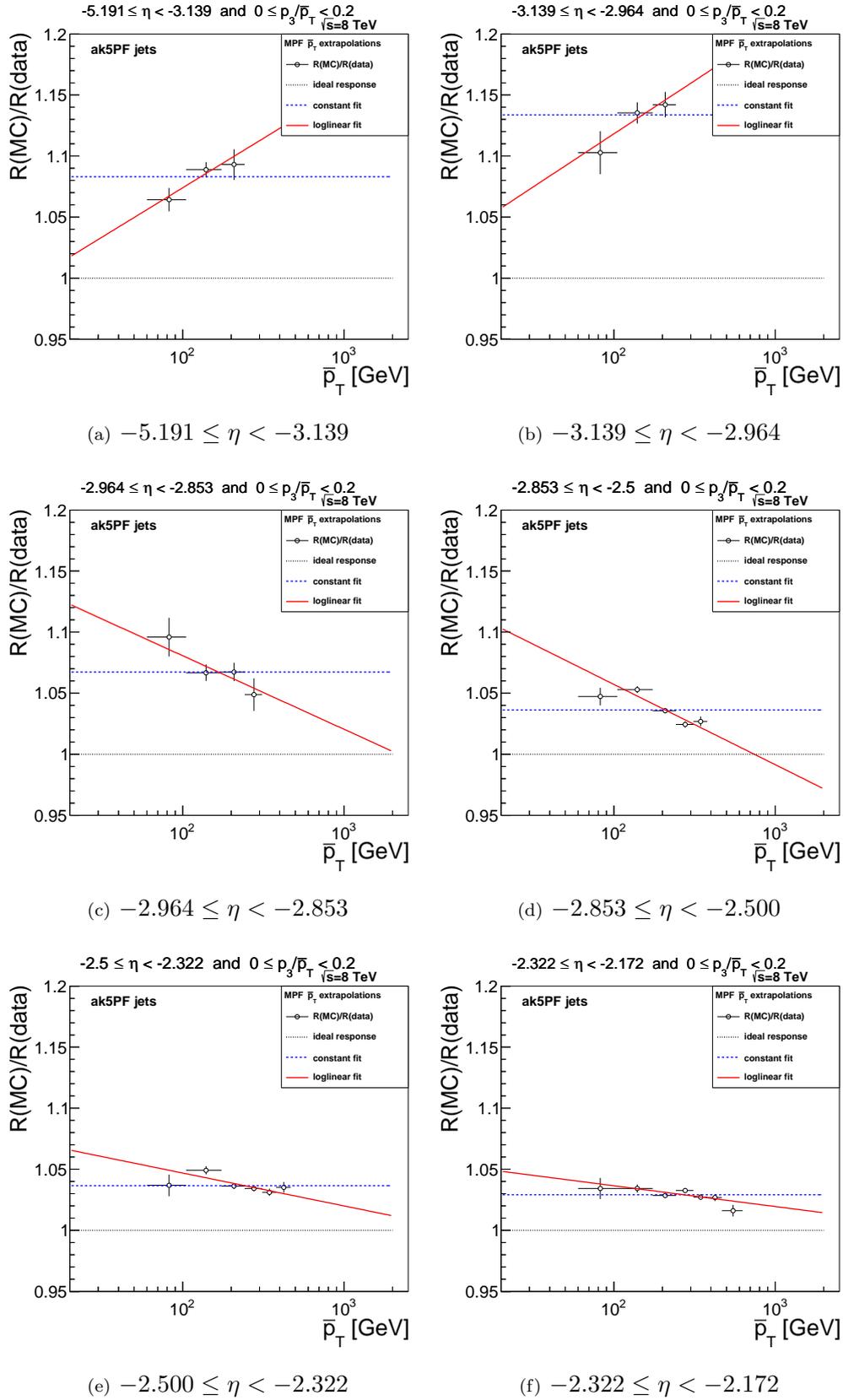
FIGURE A.16: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-5.191 \leq \eta < -2.172$ for ak5PF jets

(a) $-2.172 \leq \eta < -2.043$

(b) $-2.043 \leq \eta < -1.930$

(c) $-1.930 \leq \eta < -1.830$

(d) $-1.830 \leq \eta < -1.740$

(e) $-1.740 \leq \eta < -1.653$

(f) $-1.653 \leq \eta < -1.566$
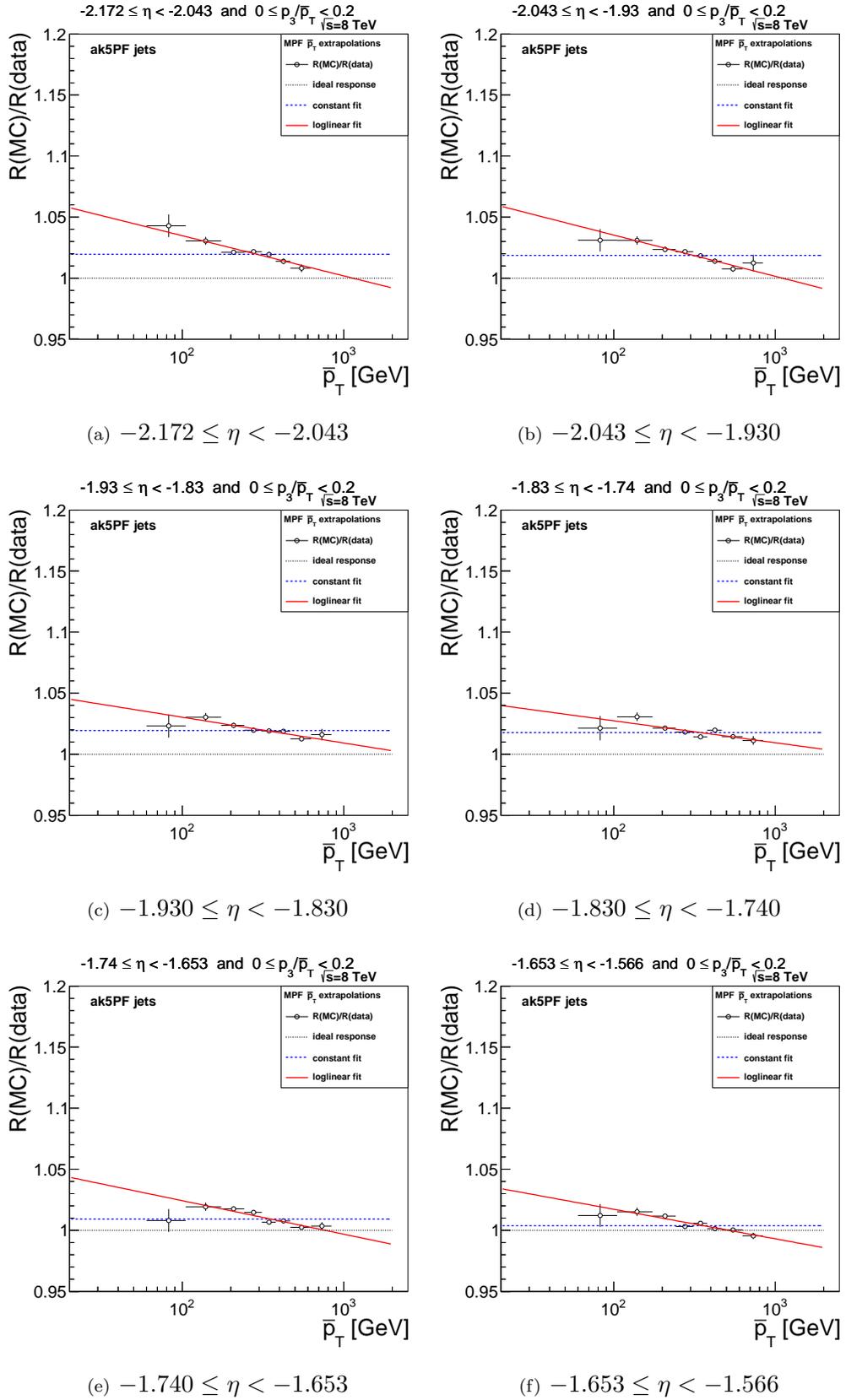
FIGURE A.17: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-2.172 \leq \eta < -1.566$ for ak5PF jets
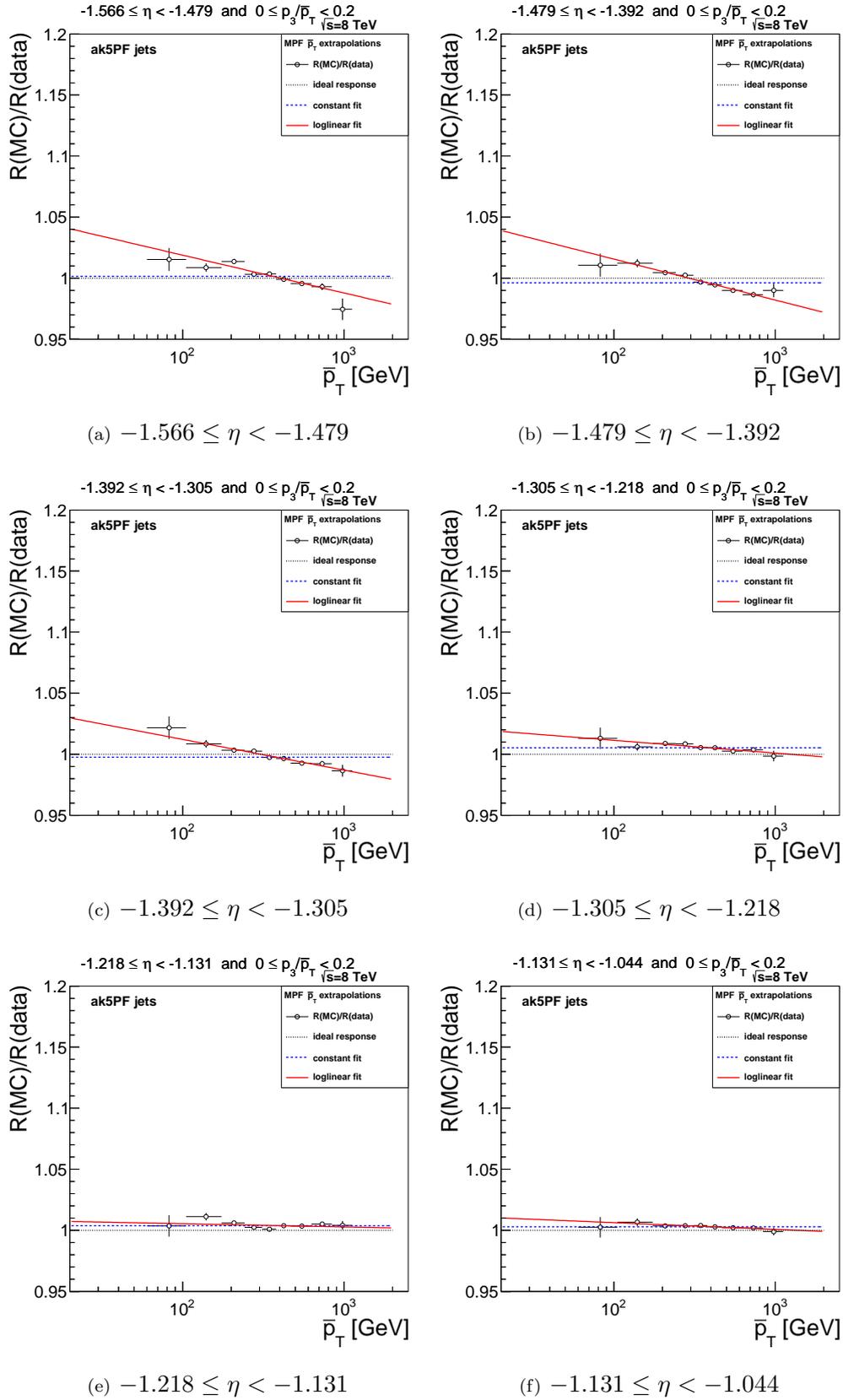
FIGURE A.18: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-1.566 \leq \eta < -1.044$ for ak5PF jets

(a) $-1.044 \leq \eta < -0.957$

(b) $-0.957 \leq \eta < -0.879$

(c) $-0.879 \leq \eta < -0.783$

(d) $-0.783 \leq \eta < -0.696$

(e) $-0.696 \leq \eta < -0.609$

(f) $-0.609 \leq \eta < -0.522$
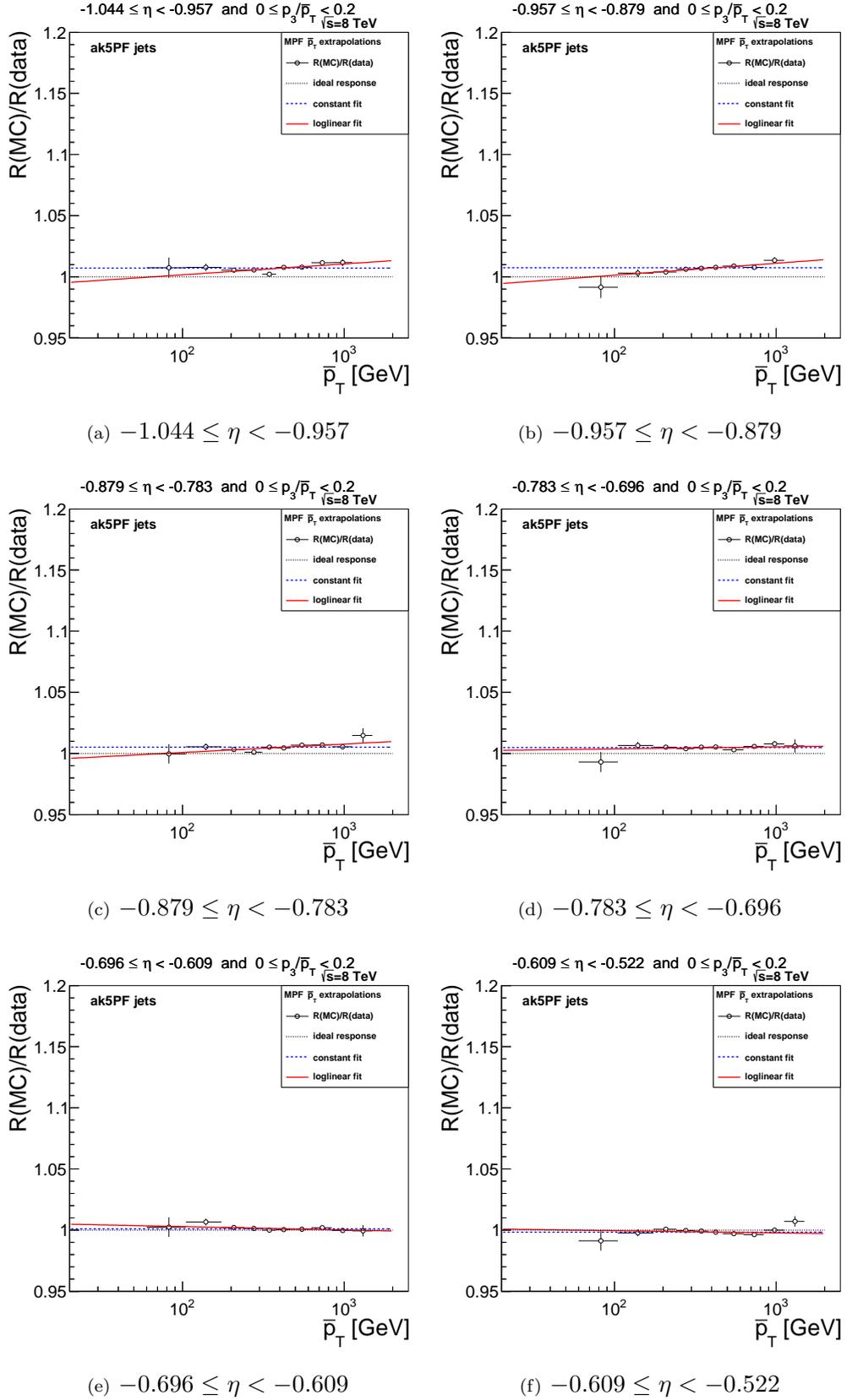
FIGURE A.19: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-1.044 \leq \eta < -0.522$ for ak5PF jets
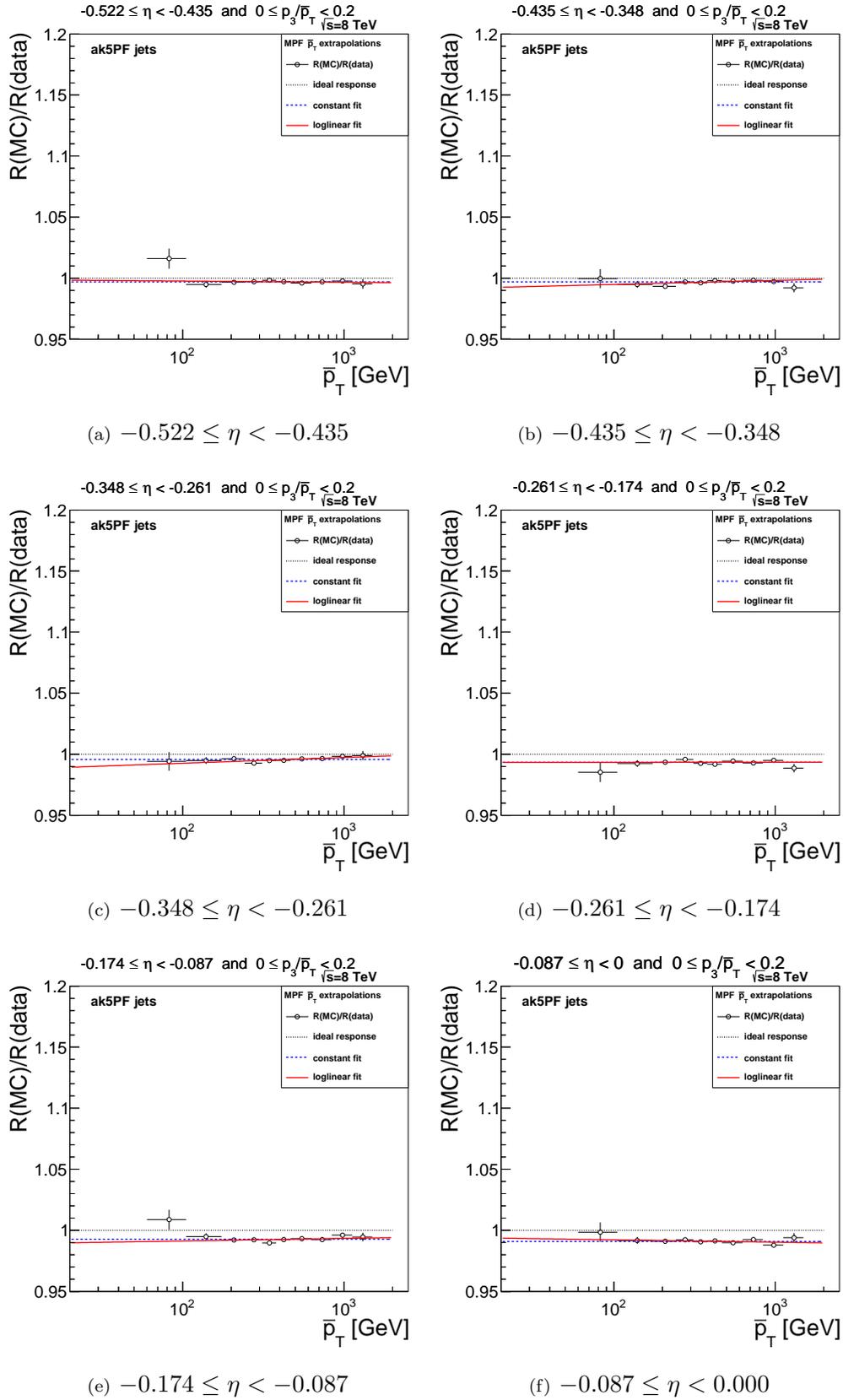
FIGURE A.20: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-0.522 \leq \eta < 0.000$ for ak5PF jets
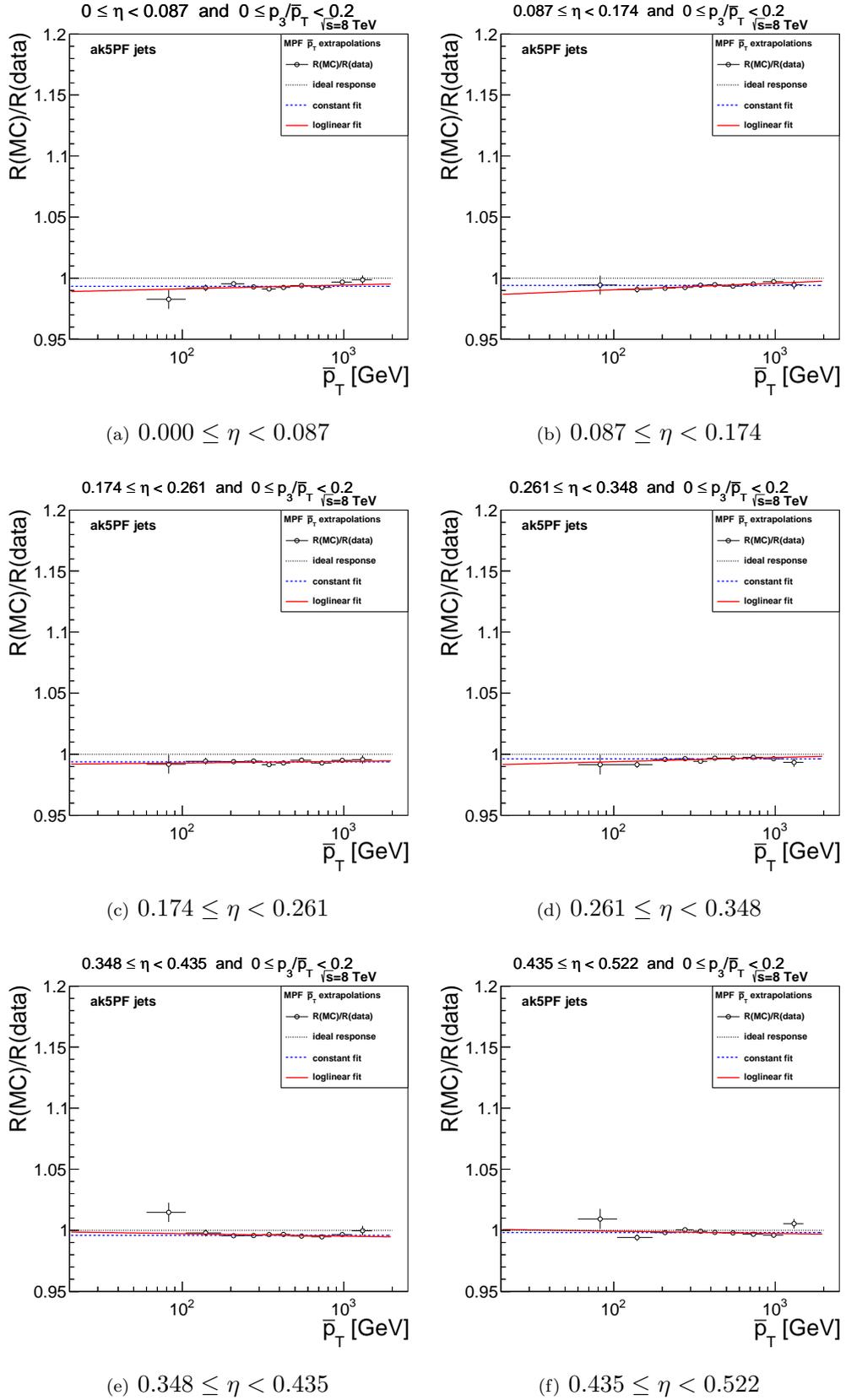
FIGURE A.21: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $0.000 \leq \eta < 0.522$ for ak5PF jets

(a) $0.522 \leq \eta < 0.609$

(b) $0.609 \leq \eta < 0.696$

(c) $0.696 \leq \eta < 0.783$

(d) $0.783 \leq \eta < 0.879$

(e) $0.879 \leq \eta < 0.957$

(f) $0.957 \leq \eta < 1.044$
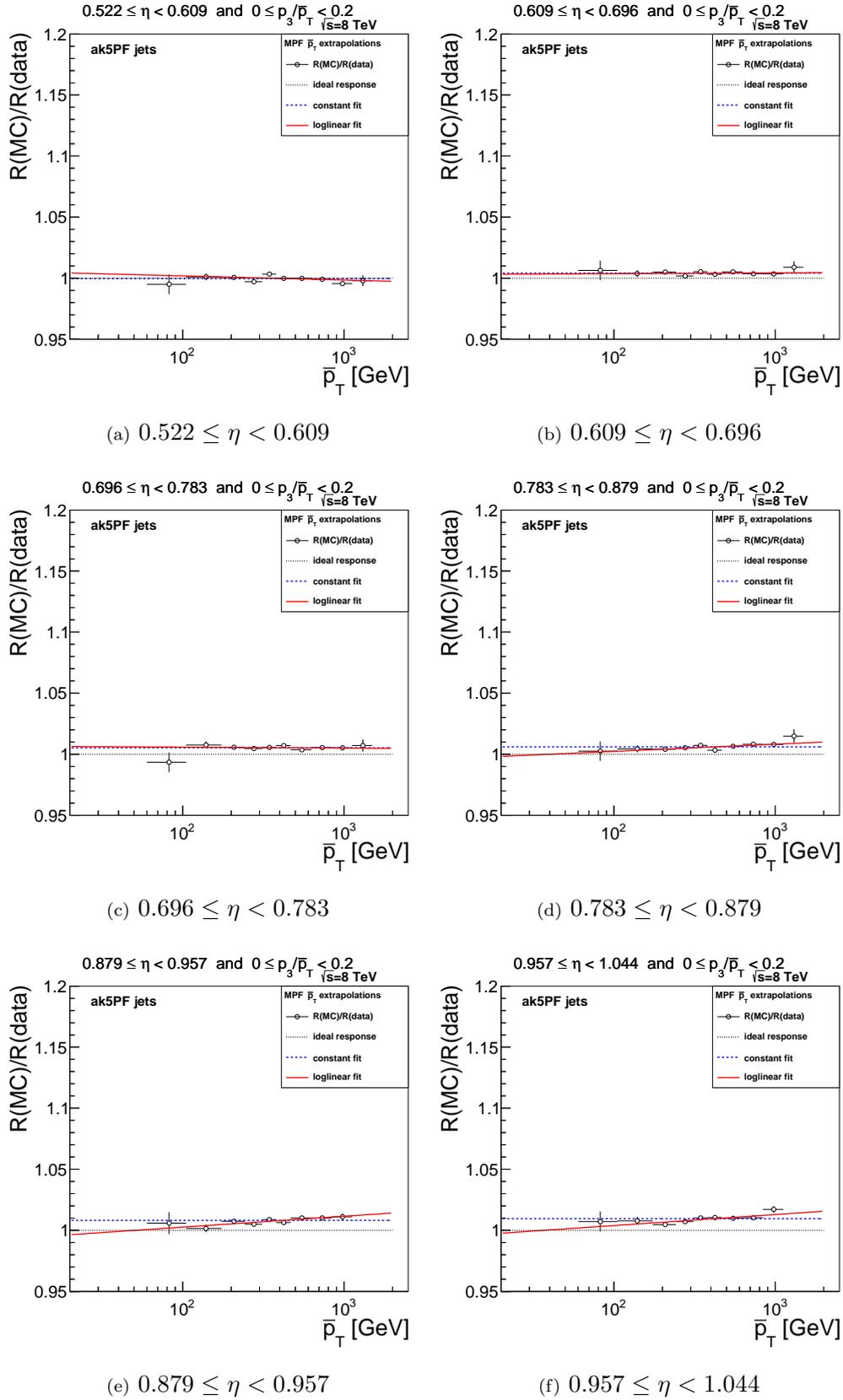
FIGURE A.22: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $0.522 \leq \eta < 1.044$ for ak5PF jets

(a) $1.044 \leq \eta < 1.131$

(b) $1.131 \leq \eta < 1.218$

(c) $1.218 \leq \eta < 1.305$

(d) $1.305 \leq \eta < 1.392$

(e) $1.392 \leq \eta < 1.479$

(f) $1.479 \leq \eta < 1.566$
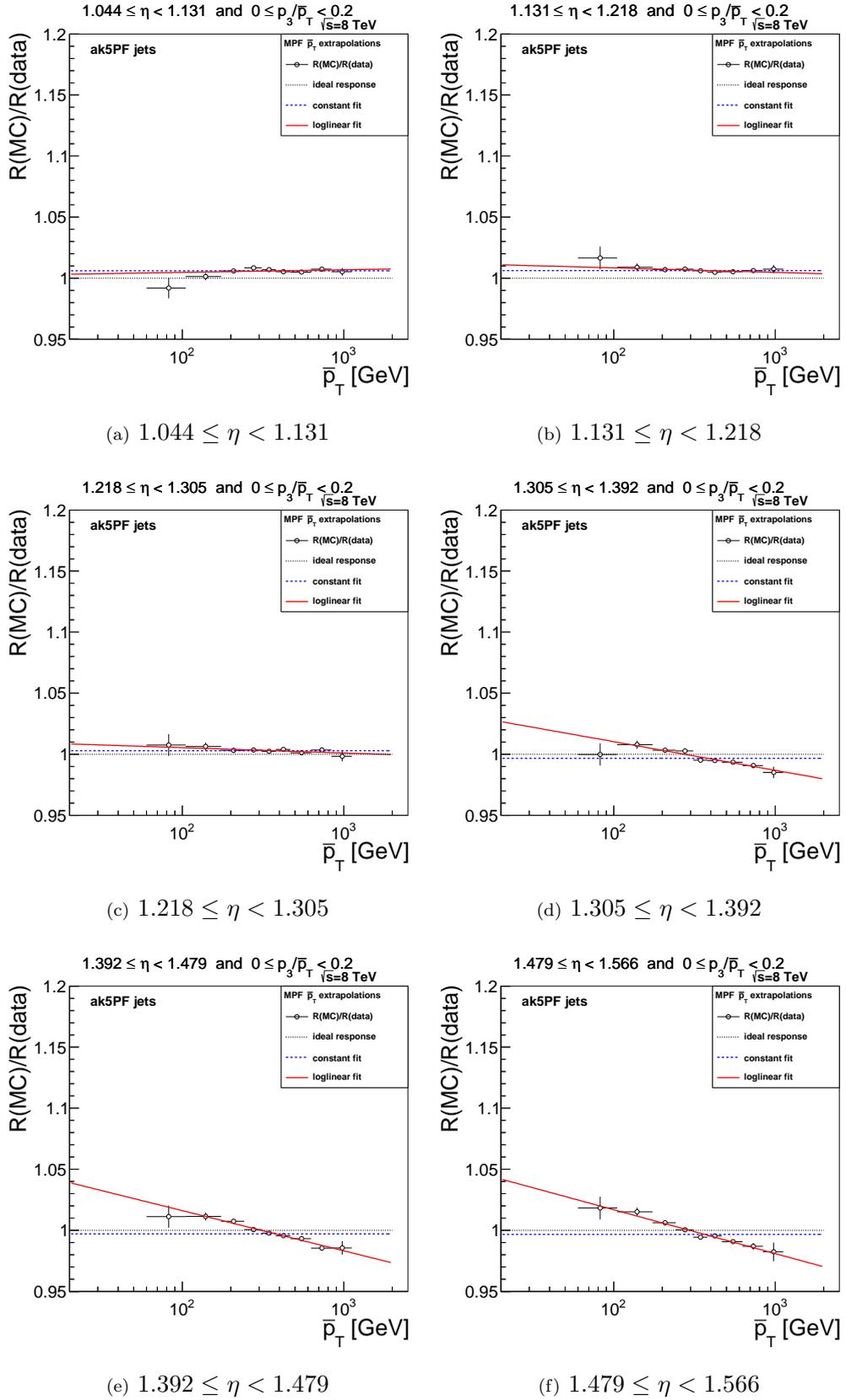
FIGURE A.23: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $1.044 \leq \eta < 1.566$ for ak5PF jets

(a) $1.566 \le \eta < 1.653$

(b) $1.653 \le \eta < 1.740$

(c) $1.740 \le \eta < 1.830$

(d) $1.830 \le \eta < 1.930$

(e) $1.930 \le \eta < 2.043$

(f) $2.043 \le \eta < 2.172$
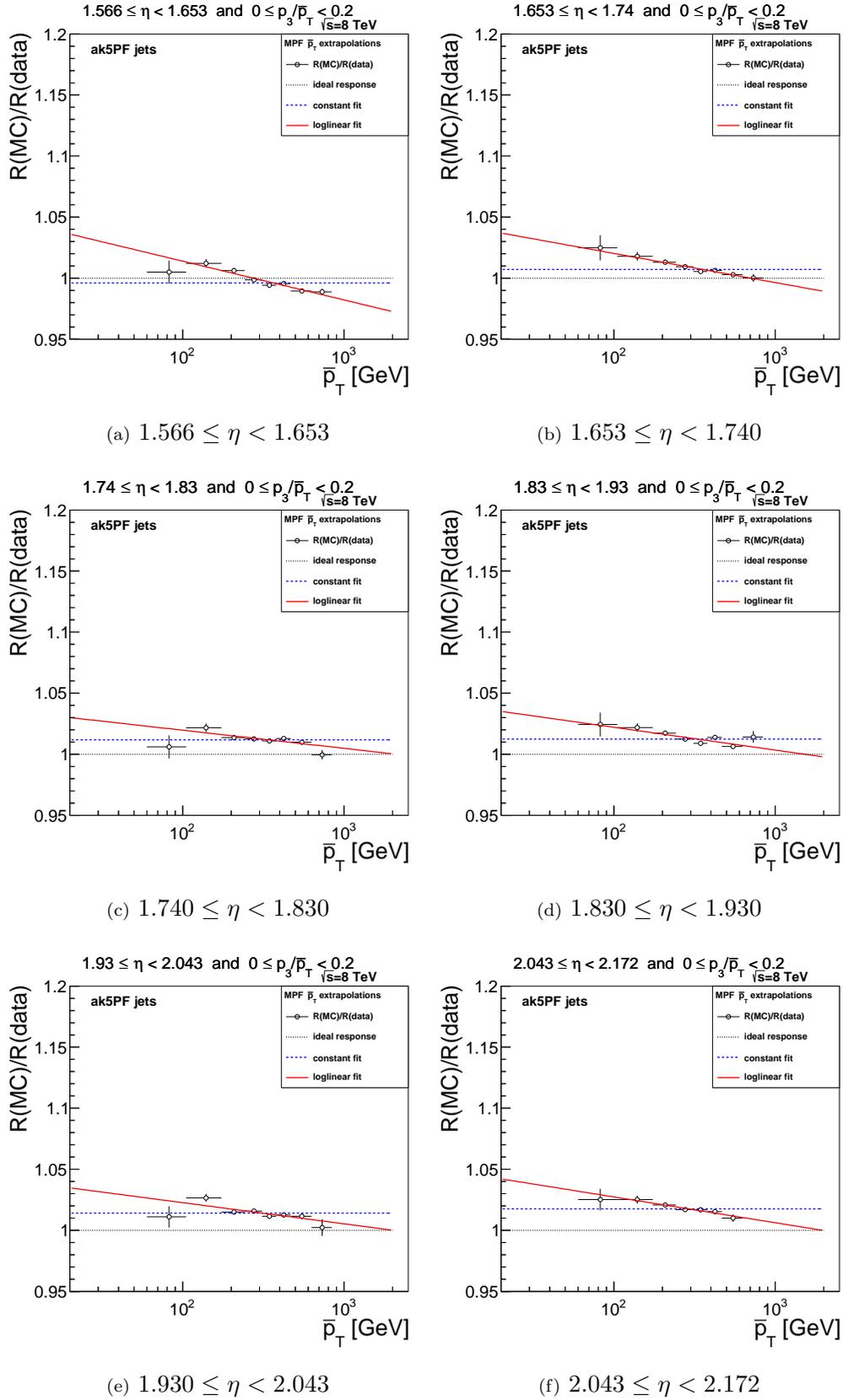
FIGURE A.24: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $1.566 \le \eta < 2.172$ for ak5PF jets

(a) $2.172 \leq \eta < 2.322$

(b) $2.322 \leq \eta < 2.500$

(c) $2.500 \leq \eta < 2.853$

(d) $2.853 \leq \eta < 2.964$

(e) $2.964 \leq \eta < 3.139$

(f) $3.139 \leq \eta < 5.191$
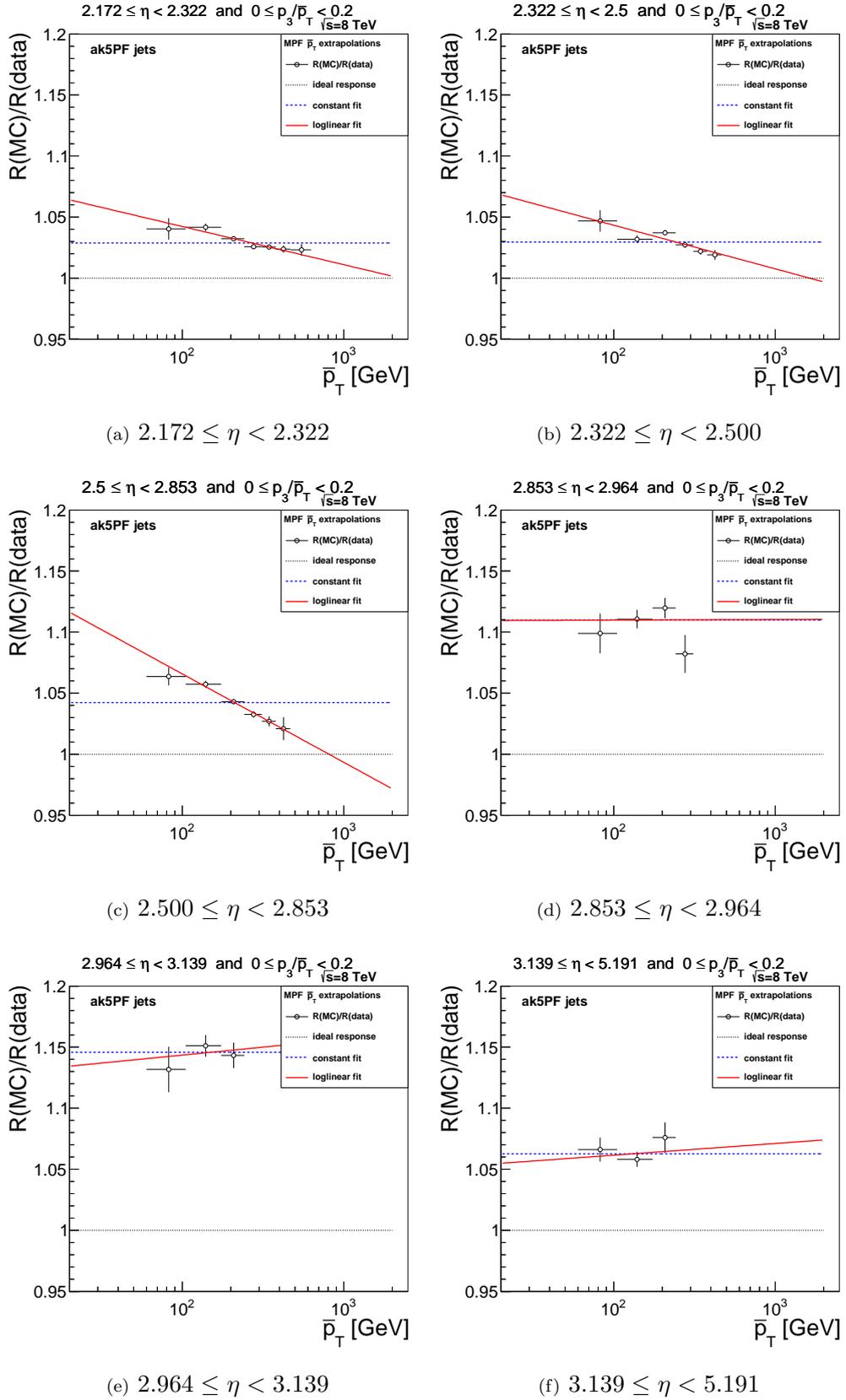
FIGURE A.25: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $2.172 \leq \eta < 5.191$ for ak5PF jets

A.3.2 $\bar{p}_T$ extrapolations - ak5PF+CHS

All extrapolations over $\bar{p}_T$ for ak5PF+CHS jets are shown in a constant and a loglinear $([0] + [1] \cdot \log(x))$ version.

(a) $-5.191 \leq \eta < -3.139$

(b) $-3.139 \leq \eta < -2.964$

(c) $-2.964 \leq \eta < -2.853$

(d) $-2.853 \leq \eta < -2.500$

(e) $-2.500 \leq \eta < -2.322$

(f) $-2.322 \leq \eta < -2.172$
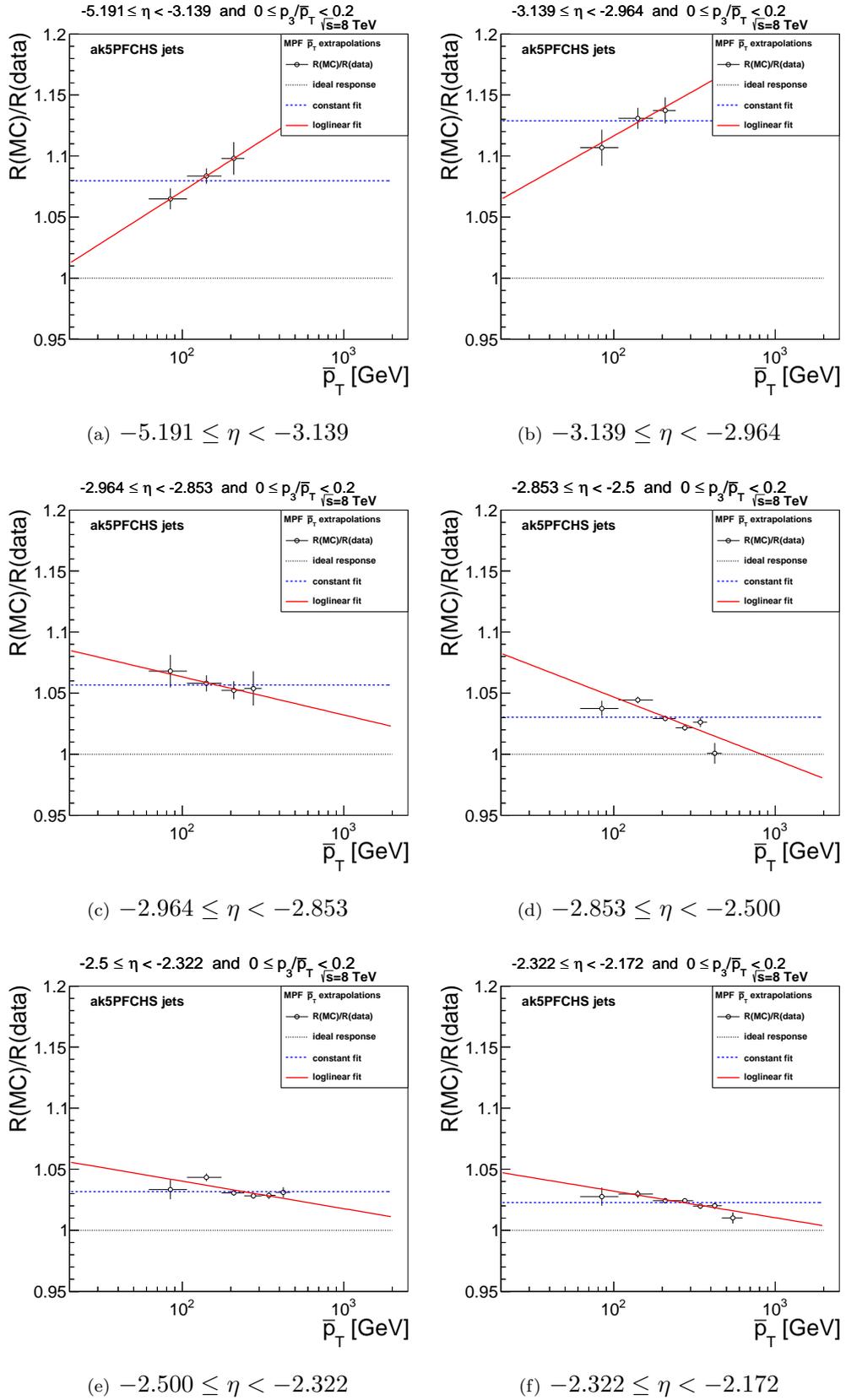
FIGURE A.26: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-5.191 \leq \eta < -2.172$ for ak5PF+CHS jets
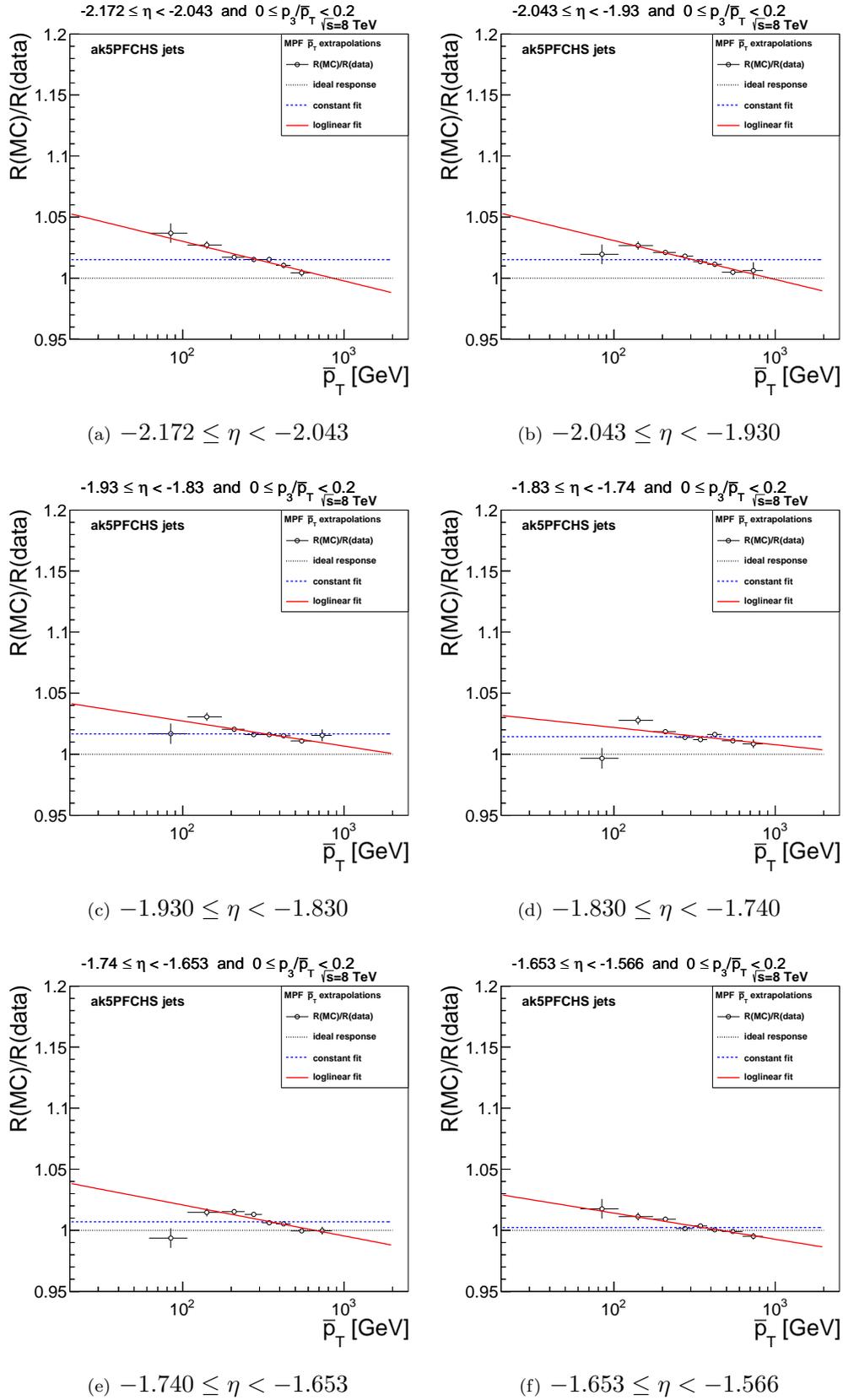
FIGURE A.27: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-2.172 \leq \eta < -1.566$ for ak5PF+CHS jets
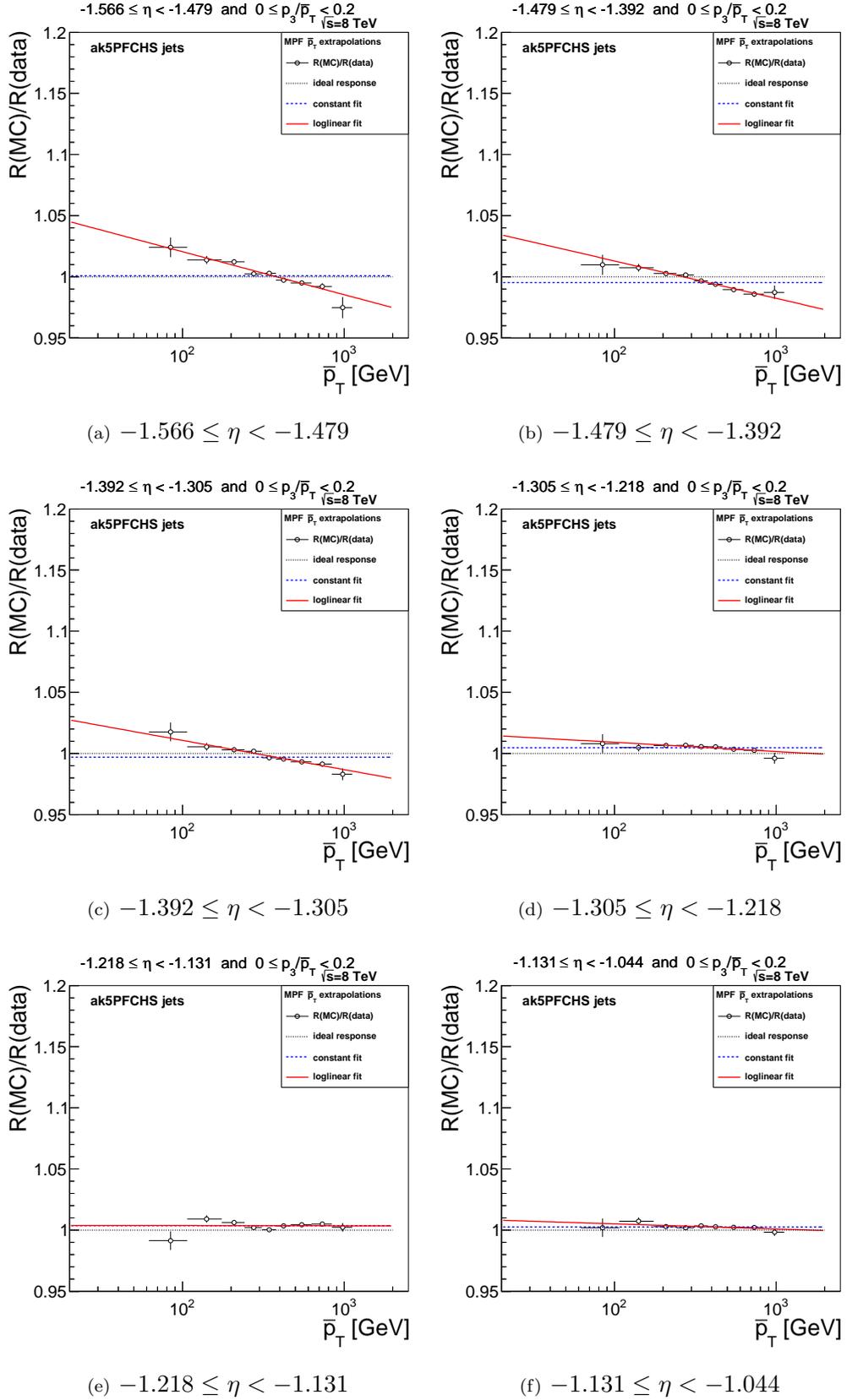
FIGURE A.28: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-1.566 \leq \eta < -1.044$ for ak5PF+CHS jets
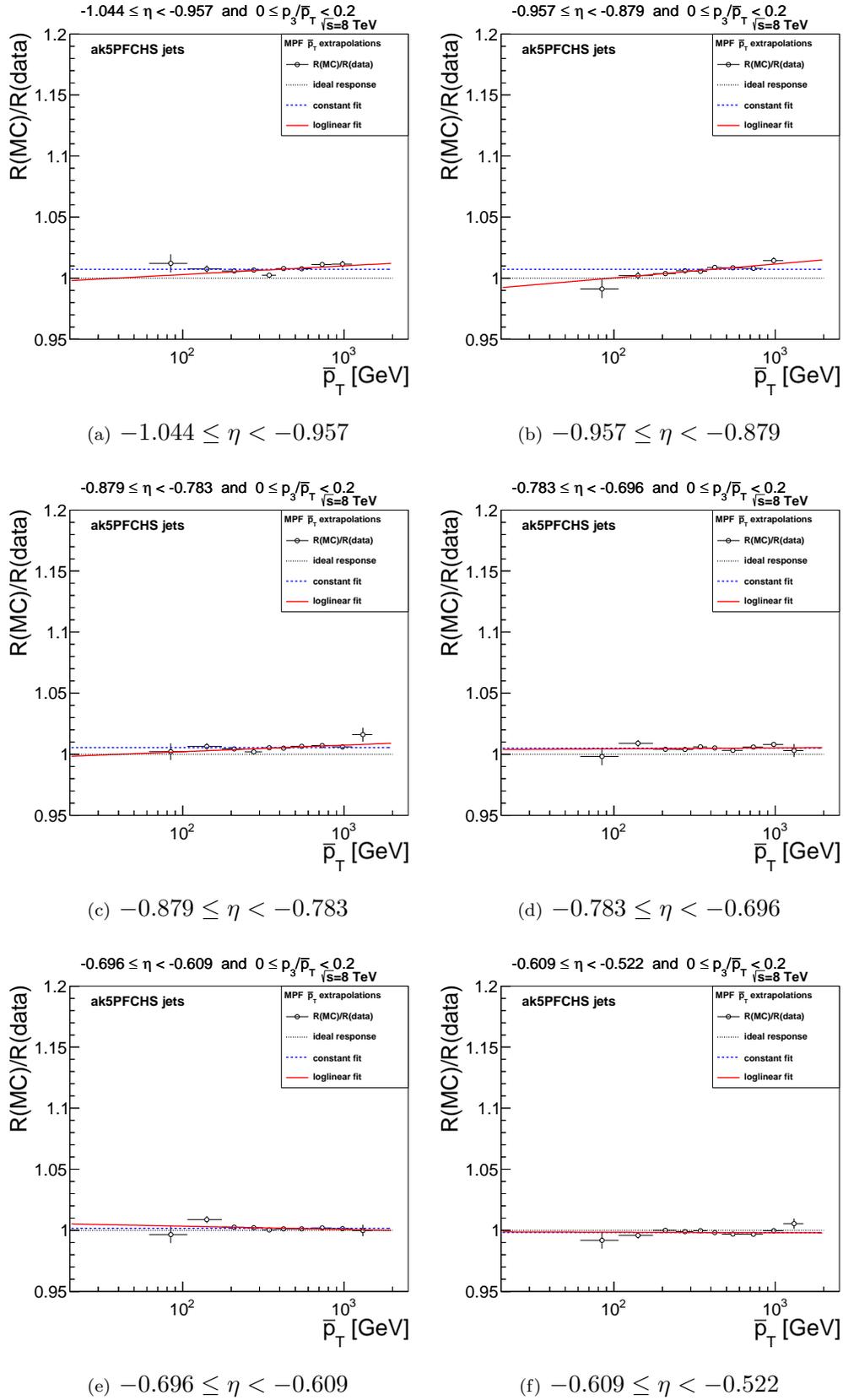
FIGURE A.29: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-1.044 \le \eta < -0.522$ for ak5PF+CHS jets

(a) $-0.522 \leq \eta < -0.435$

(b) $-0.435 \leq \eta < -0.348$

(c) $-0.348 \leq \eta < -0.261$

(d) $-0.261 \leq \eta < -0.174$

(e) $-0.174 \leq \eta < -0.087$

(f) $-0.087 \leq \eta < 0.000$
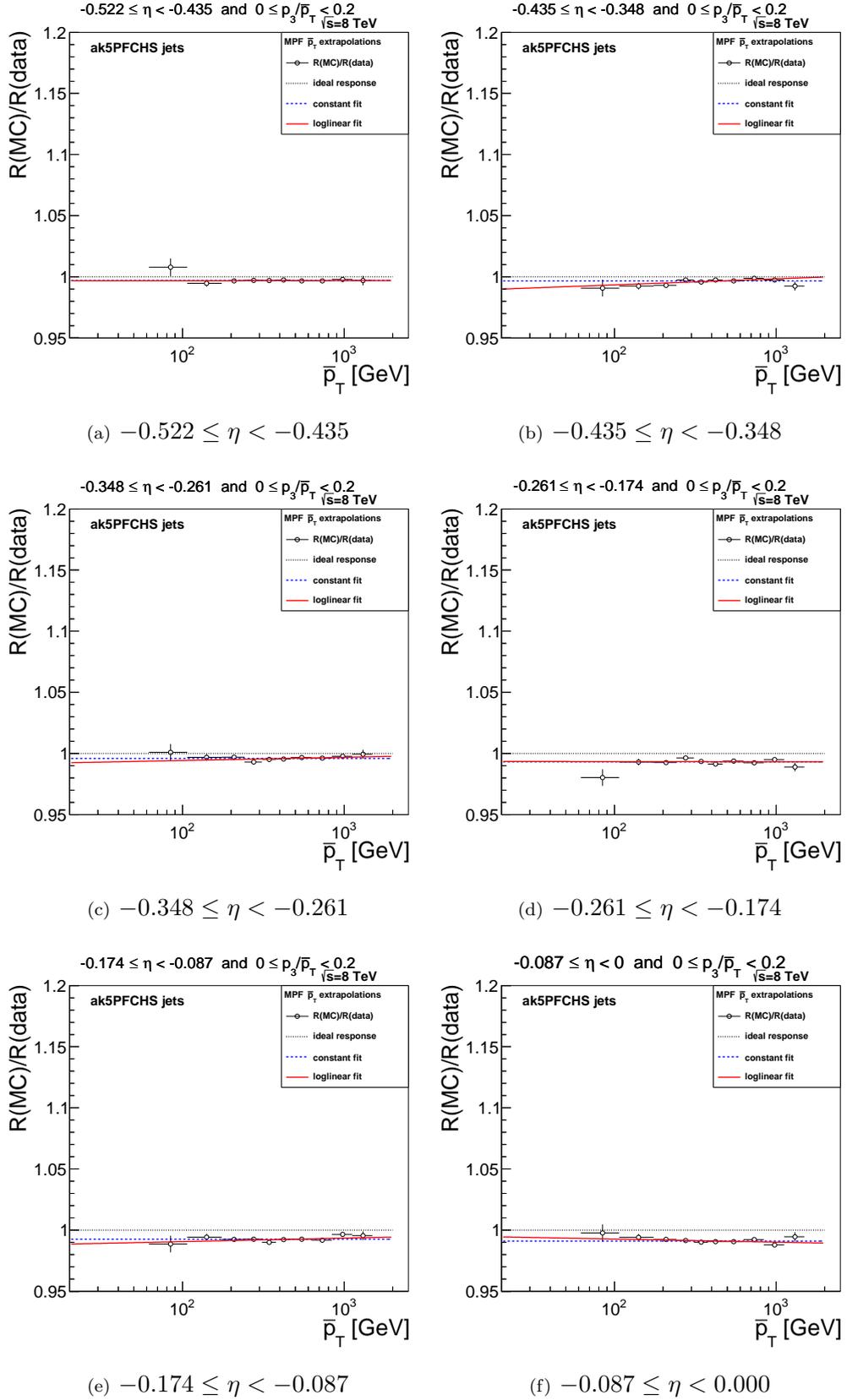
FIGURE A.30: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-0.522 \leq \eta < 0.000$ for ak5PF+CHS jets
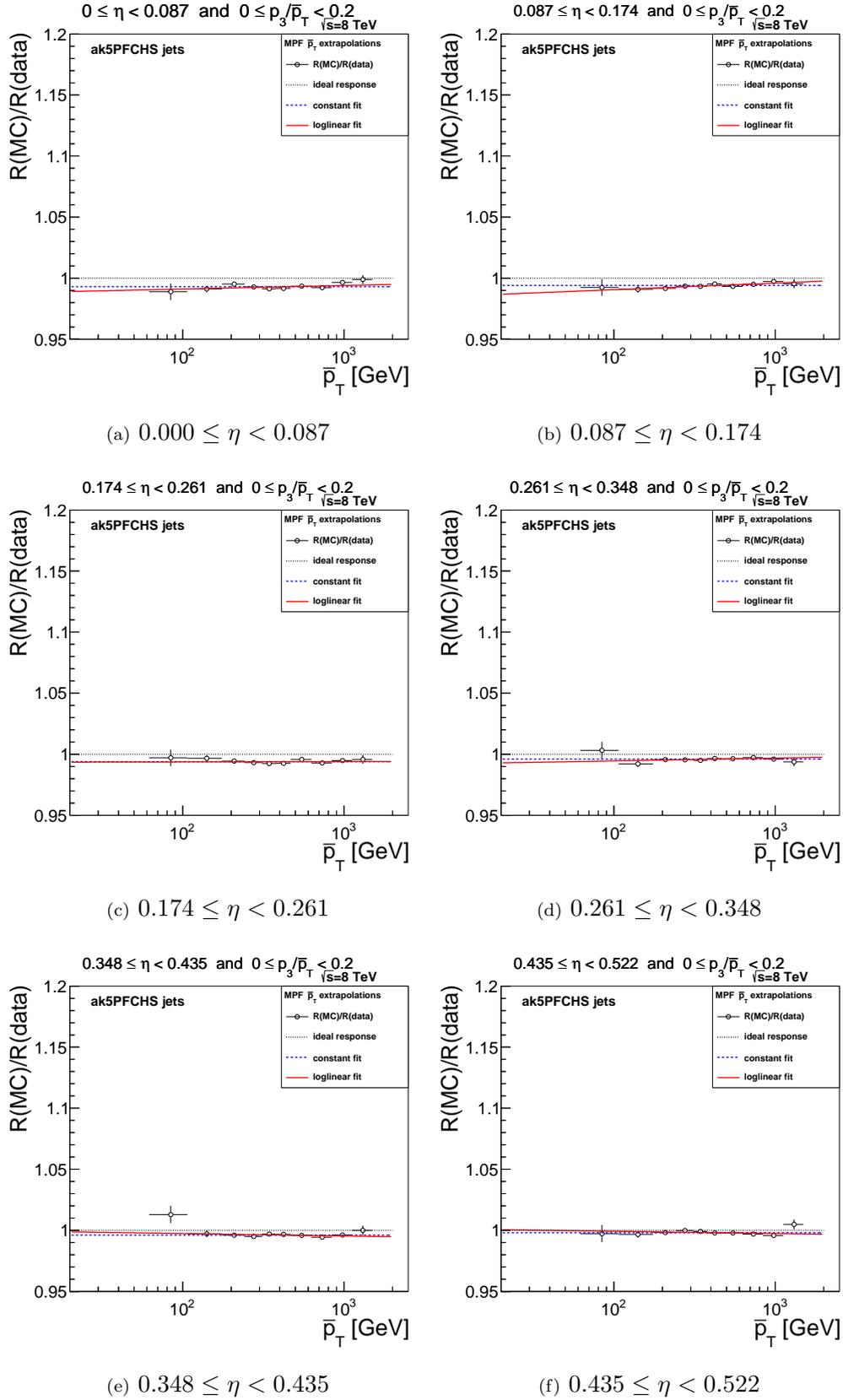
FIGURE A.31: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $0.000 \leq \eta < 0.522$ for ak5PF+CHS jets

(a) $0.522 \leq \eta < 0.609$

(b) $0.609 \leq \eta < 0.696$

(c) $0.696 \leq \eta < 0.783$

(d) $0.783 \leq \eta < 0.879$

(e) $0.879 \leq \eta < 0.957$

(f) $0.957 \leq \eta < 1.044$
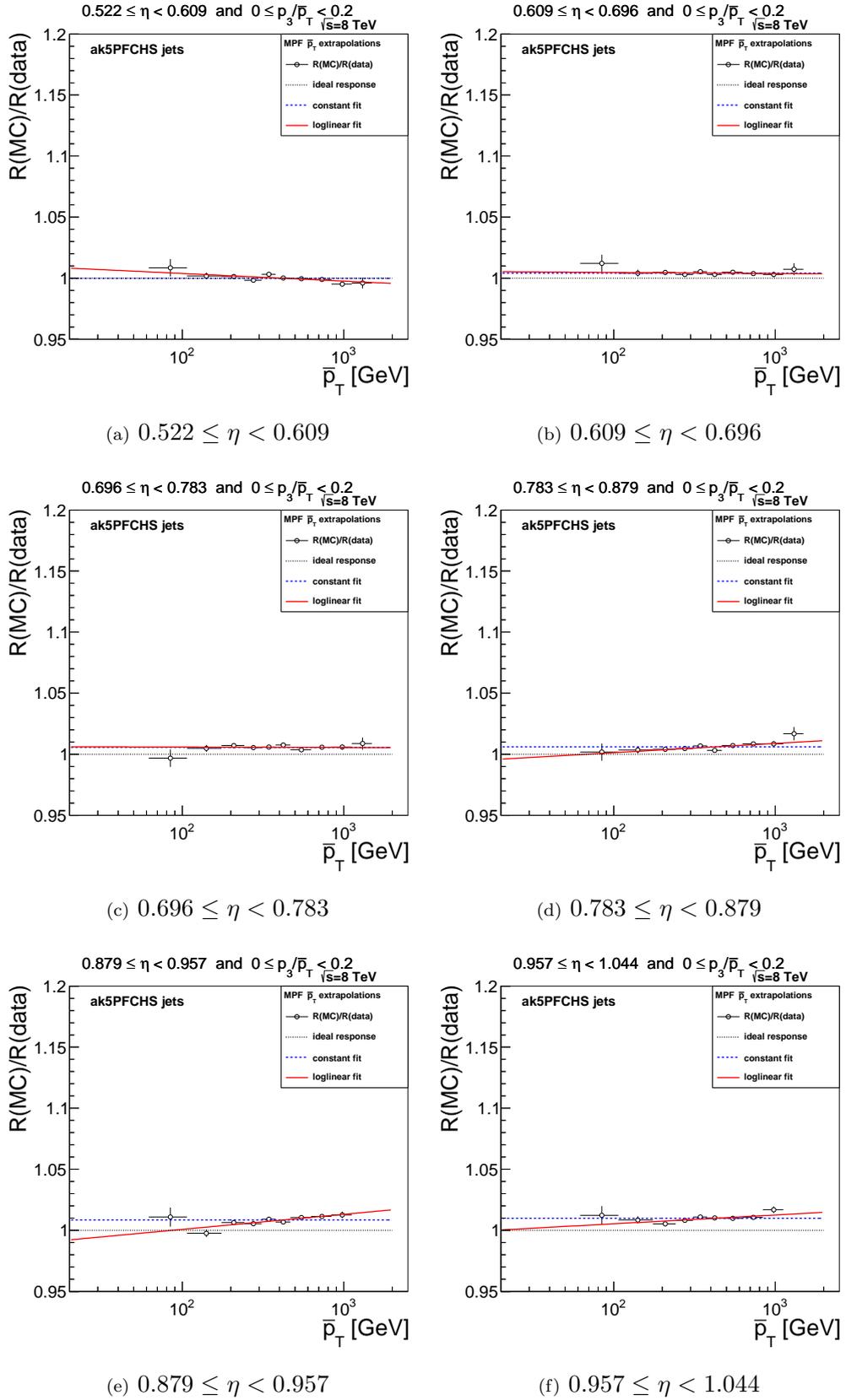
FIGURE A.32: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $0.522 \leq \eta < 1.044$ for ak5PF+CHS jets

(a) $1.044 \leq \eta < 1.131$

(b) $1.131 \leq \eta < 1.218$

(c) $1.218 \leq \eta < 1.305$

(d) $1.305 \leq \eta < 1.392$

(e) $1.392 \leq \eta < 1.479$

(f) $1.479 \leq \eta < 1.566$
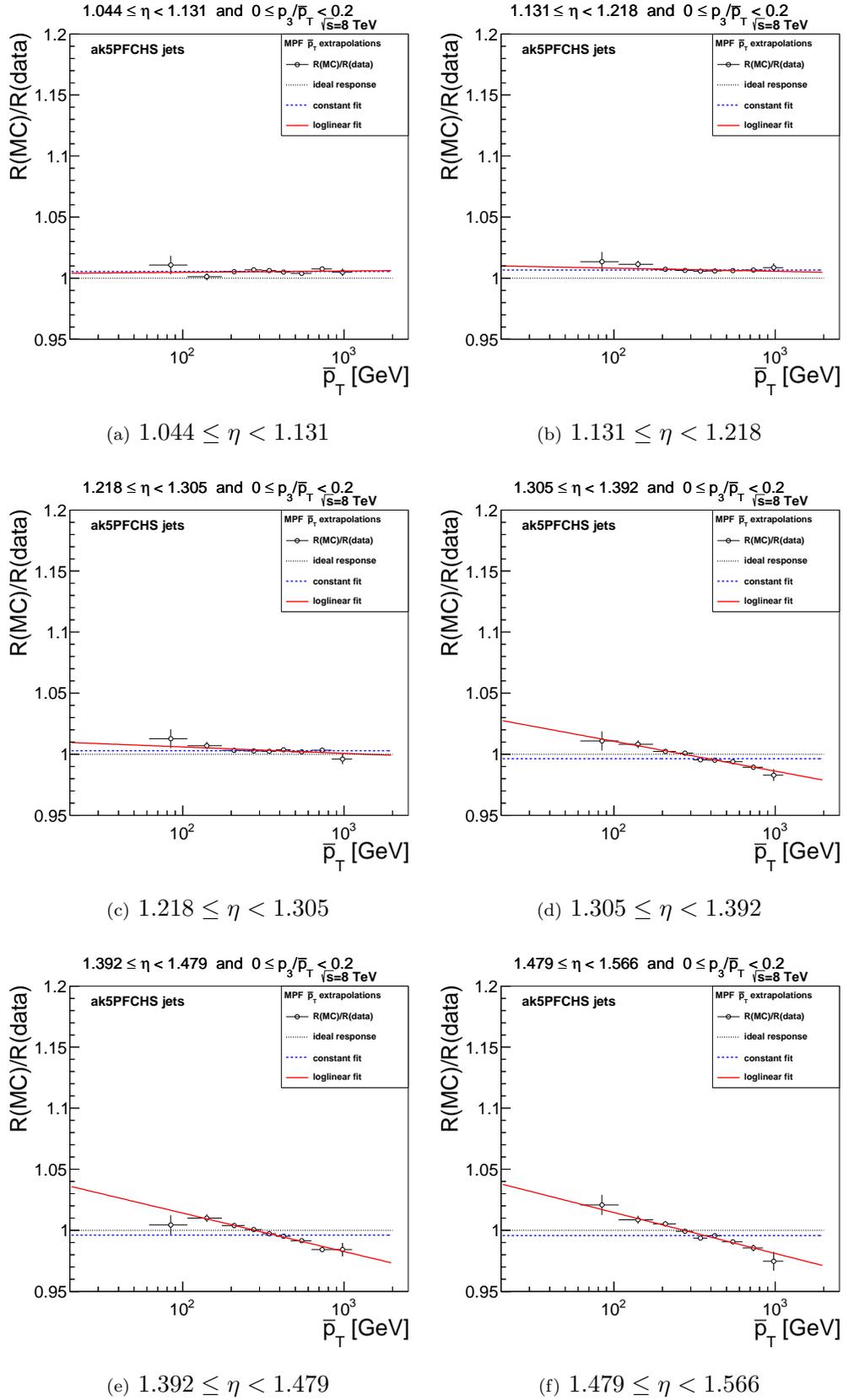
FIGURE A.33: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $1.044 \leq \eta < 1.566$ for ak5PF+CHS jets

(a) $1.566 \leq \eta < 1.653$

(b) $1.653 \leq \eta < 1.740$

(c) $1.740 \leq \eta < 1.830$

(d) $1.830 \leq \eta < 1.930$

(e) $1.930 \leq \eta < 2.043$

(f) $2.043 \leq \eta < 2.172$
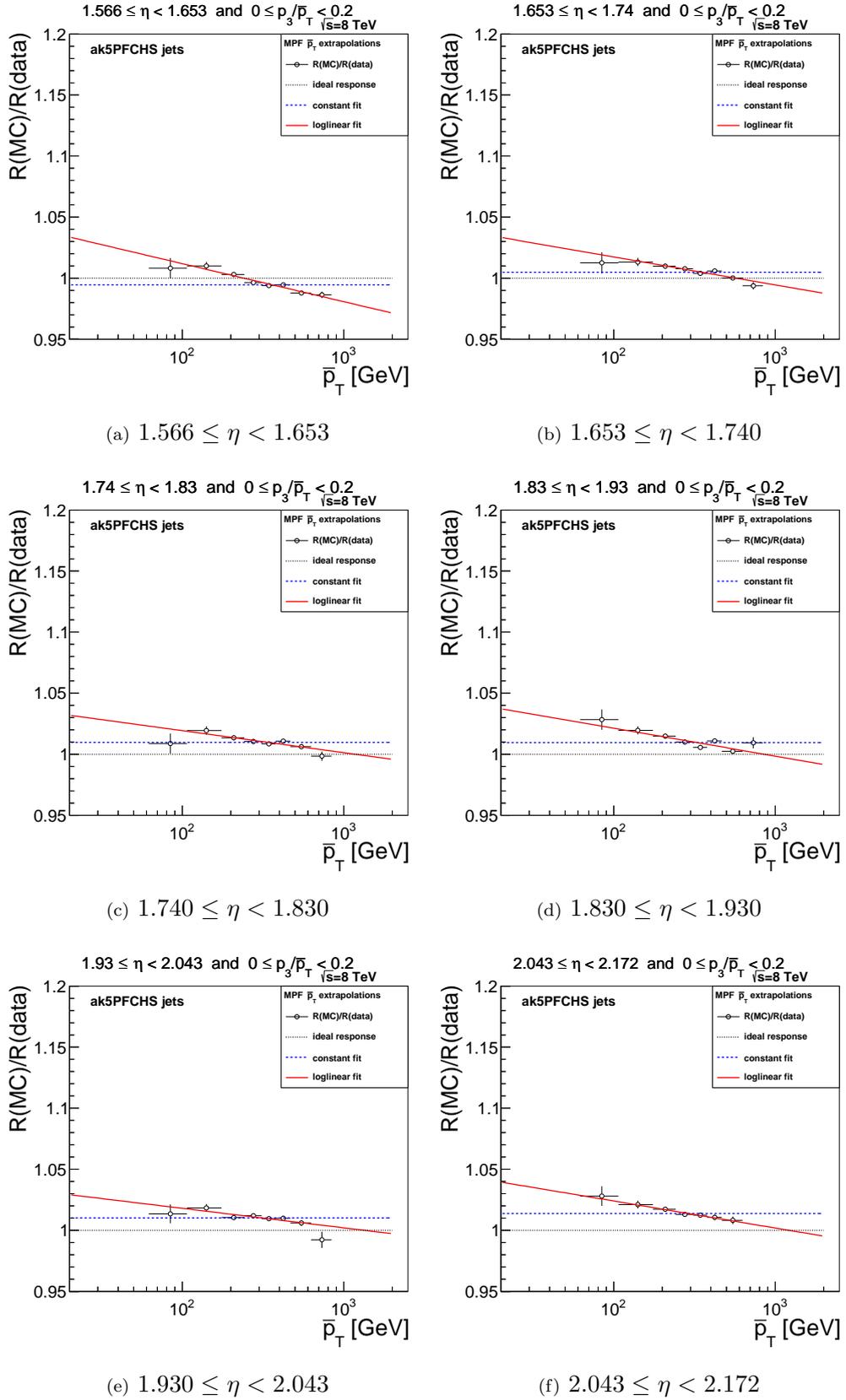
FIGURE A.34: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $1.566 \leq \eta < 2.172$ for ak5PF+CHS jets

(a) $2.172 \leq \eta < 2.322$

(b) $2.322 \leq \eta < 2.500$

(c) $2.500 \leq \eta < 2.853$

(d) $2.853 \leq \eta < 2.964$

(e) $2.964 \leq \eta < 3.139$

(f) $3.139 \leq \eta < 5.191$
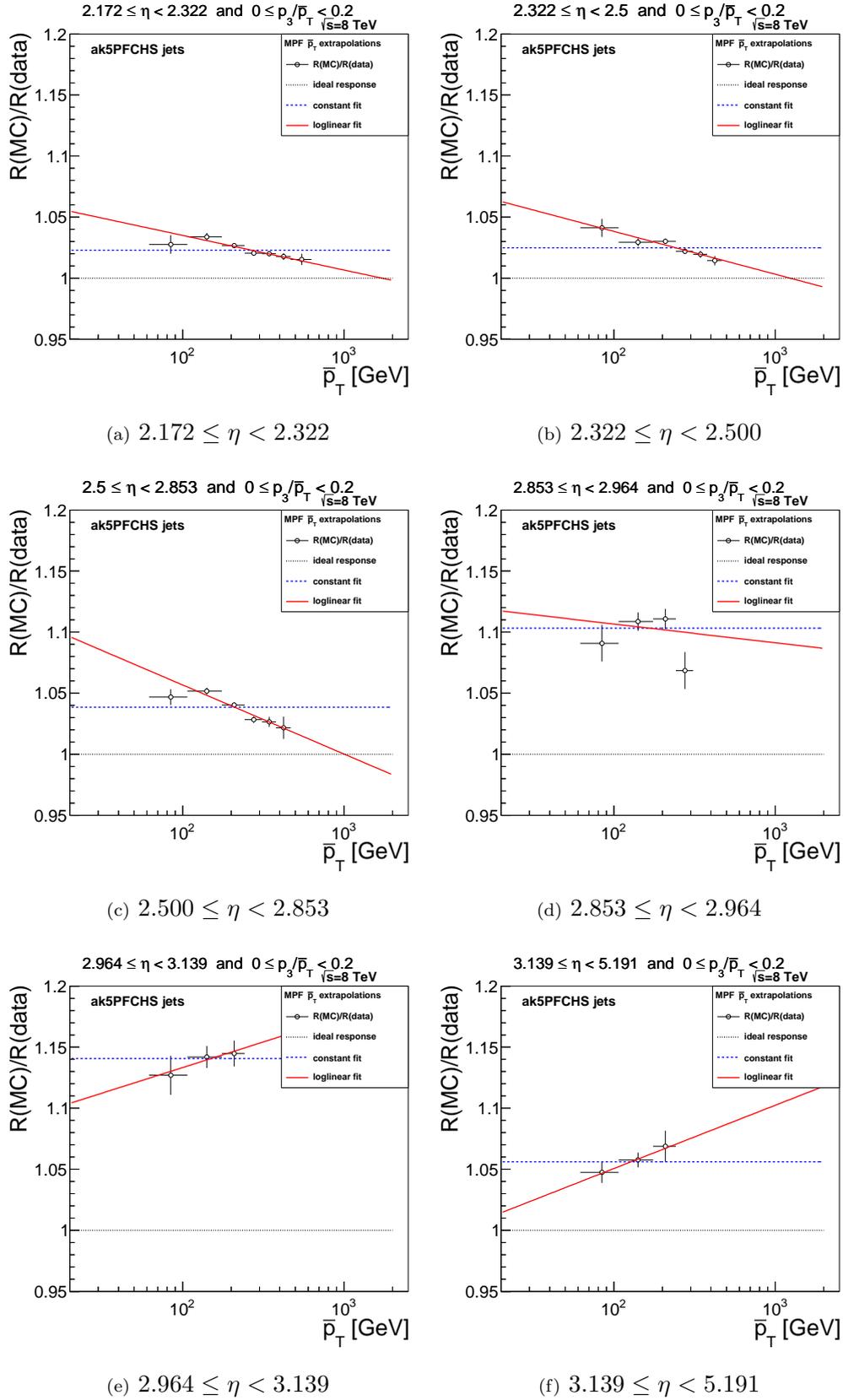
FIGURE A.35: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $2.172 \leq \eta < 5.191$ for ak5PF+CHS jets

A.3.3 $\bar{p}_T$ extrapolations - ak7PF

All extrapolations over $\bar{p}_T$ for ak7PF jets are shown in a constant and a loglinear ($[0] + [1] \cdot \log(x)$) version.
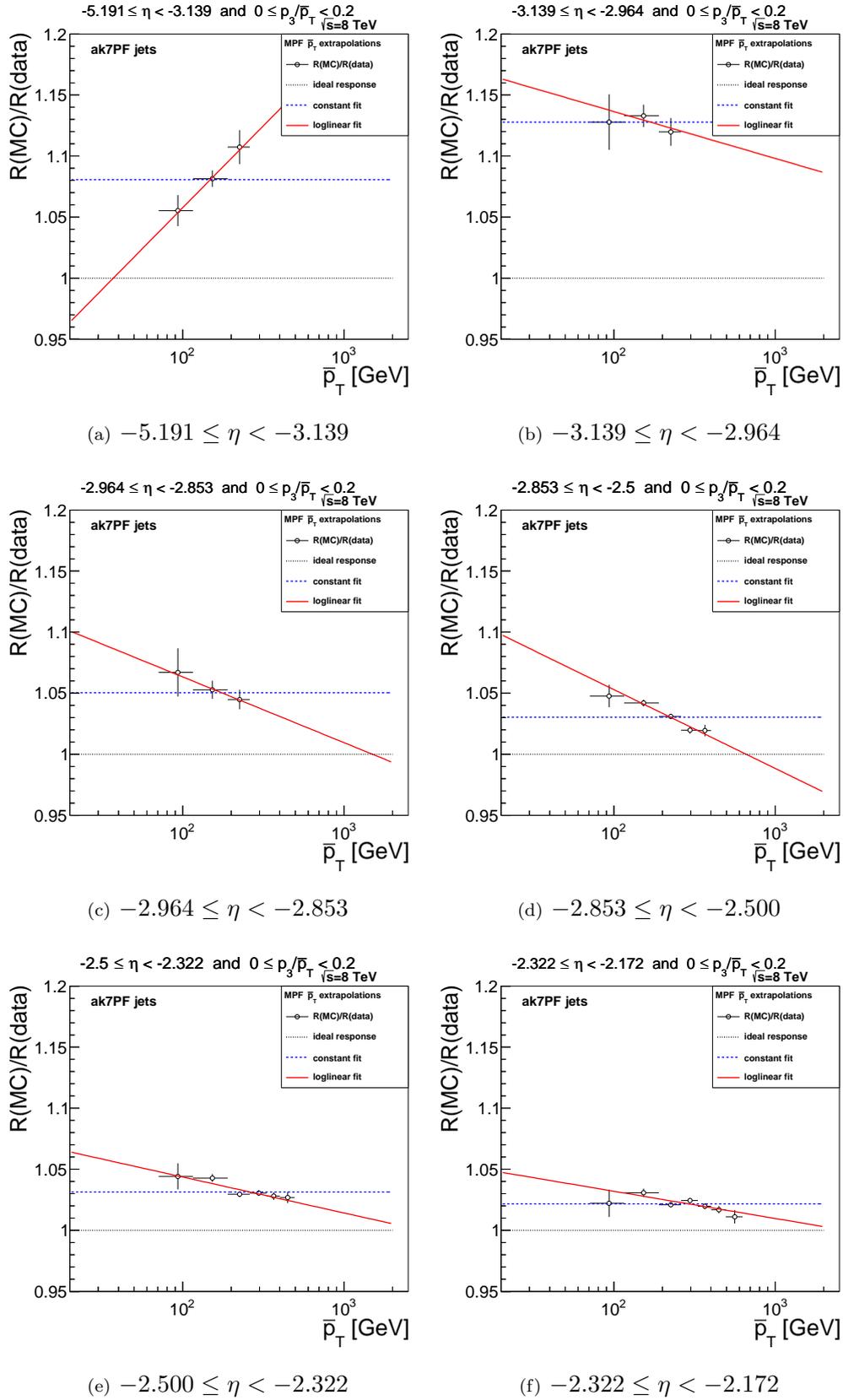
FIGURE A.36: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-5.191 \leq \eta < -2.172$ for ak7PF jets

(a) $-2.172 \leq \eta < -2.043$

(b) $-2.043 \leq \eta < -1.930$

(c) $-1.930 \leq \eta < -1.830$

(d) $-1.830 \leq \eta < -1.740$

(e) $-1.740 \leq \eta < -1.653$

(f) $-1.653 \leq \eta < -1.566$
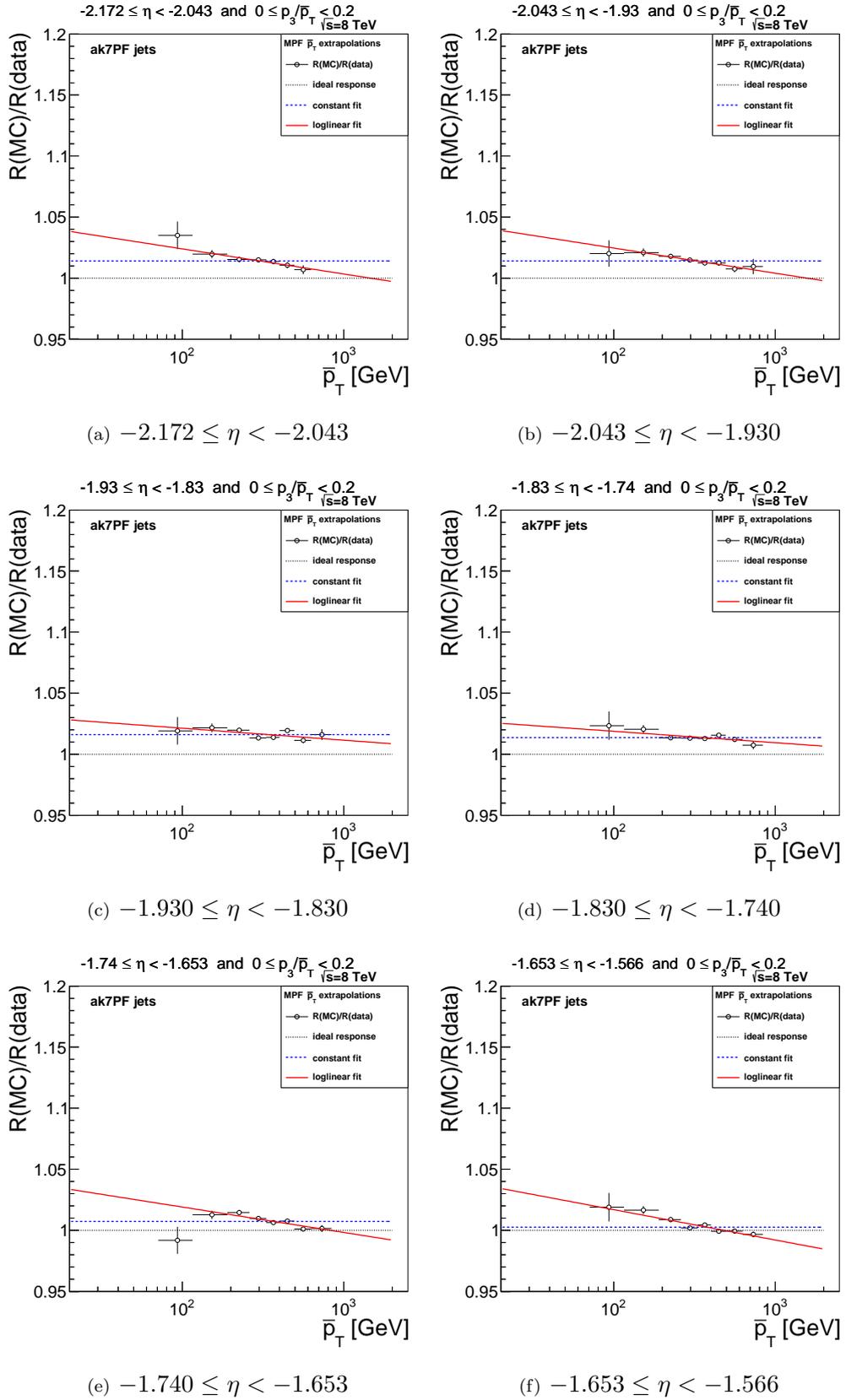
FIGURE A.37: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-2.172 \leq \eta < -1.566$ for ak7PF jets

(a) $-1.566 \leq \eta < -1.479$

(b) $-1.479 \leq \eta < -1.392$

(c) $-1.392 \leq \eta < -1.305$

(d) $-1.305 \leq \eta < -1.218$

(e) $-1.218 \leq \eta < -1.131$

(f) $-1.131 \leq \eta < -1.044$
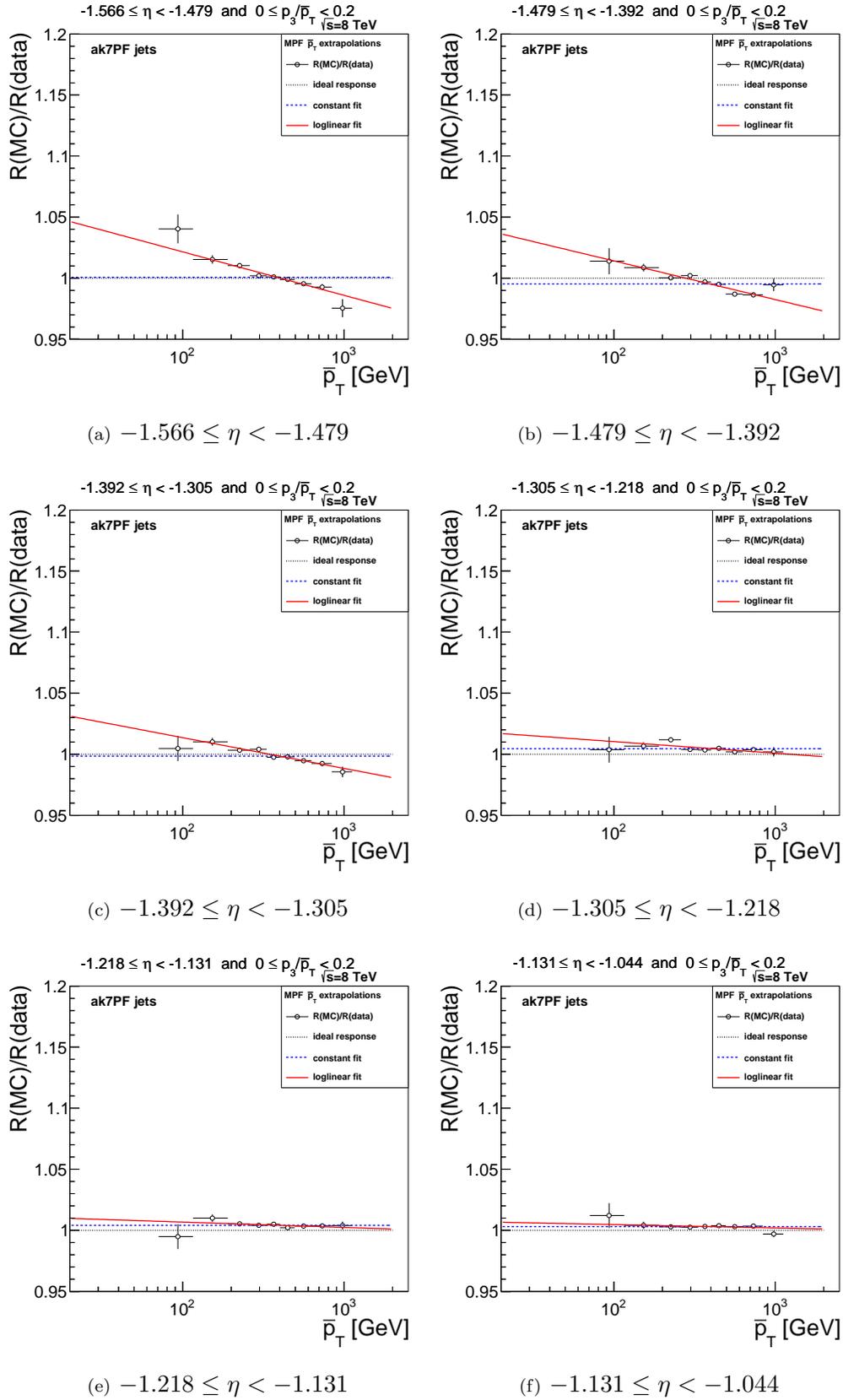
FIGURE A.38: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-1.566 \leq \eta < -1.044$ for ak7PF jets

(a) $-1.044 \leq \eta < -0.957$

(b) $-0.957 \leq \eta < -0.879$

(c) $-0.879 \leq \eta < -0.783$

(d) $-0.783 \leq \eta < -0.696$

(e) $-0.696 \leq \eta < -0.609$

(f) $-0.609 \leq \eta < -0.522$
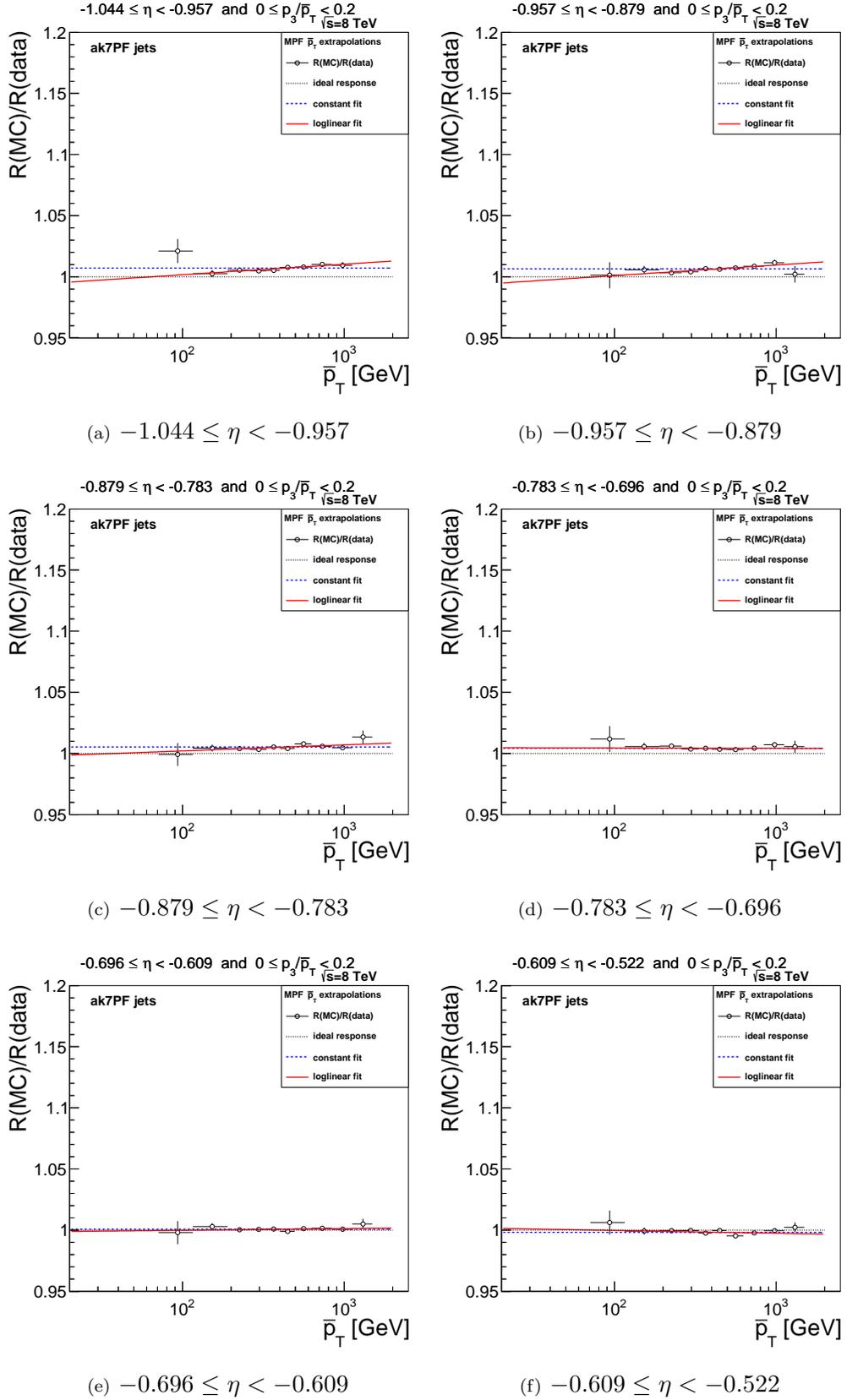
FIGURE A.39: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $-1.044 \leq \eta < -0.522$ for ak7PF jets

(a) $-0.522 \leq \eta < -0.435$

(b) $-0.435 \leq \eta < -0.348$

(c) $-0.348 \leq \eta < -0.261$

(d) $-0.261 \leq \eta < -0.174$

(e) $-0.174 \leq \eta < -0.087$

(f) $-0.087 \leq \eta < 0.000$
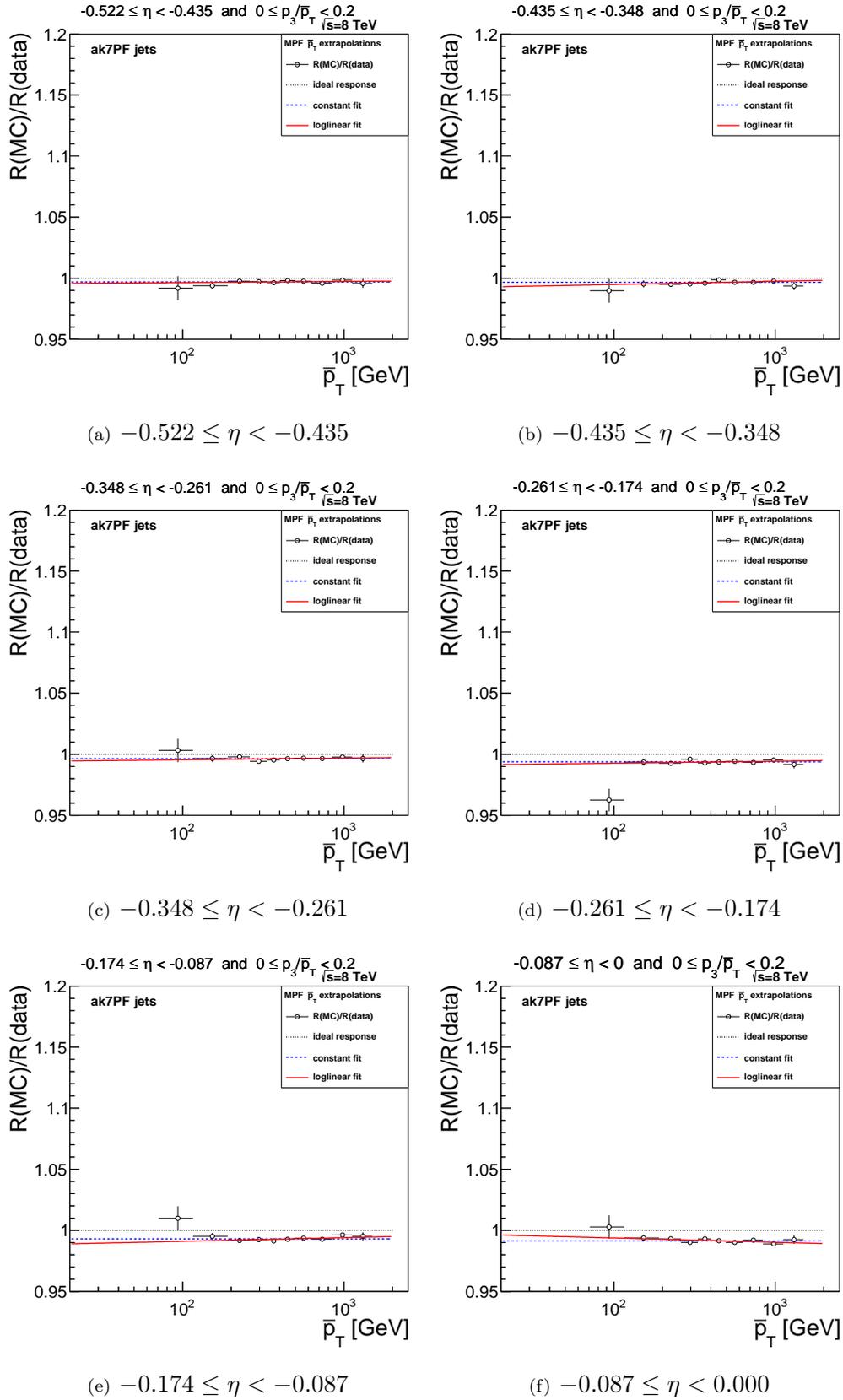
FIGURE A.40: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $-0.522 \leq \eta < 0.000$ for ak7PF jets

(a) $0.000 \leq \eta < 0.087$

(b) $0.087 \leq \eta < 0.174$

(c) $0.174 \leq \eta < 0.261$

(d) $0.261 \leq \eta < 0.348$

(e) $0.348 \leq \eta < 0.435$

(f) $0.435 \leq \eta < 0.522$
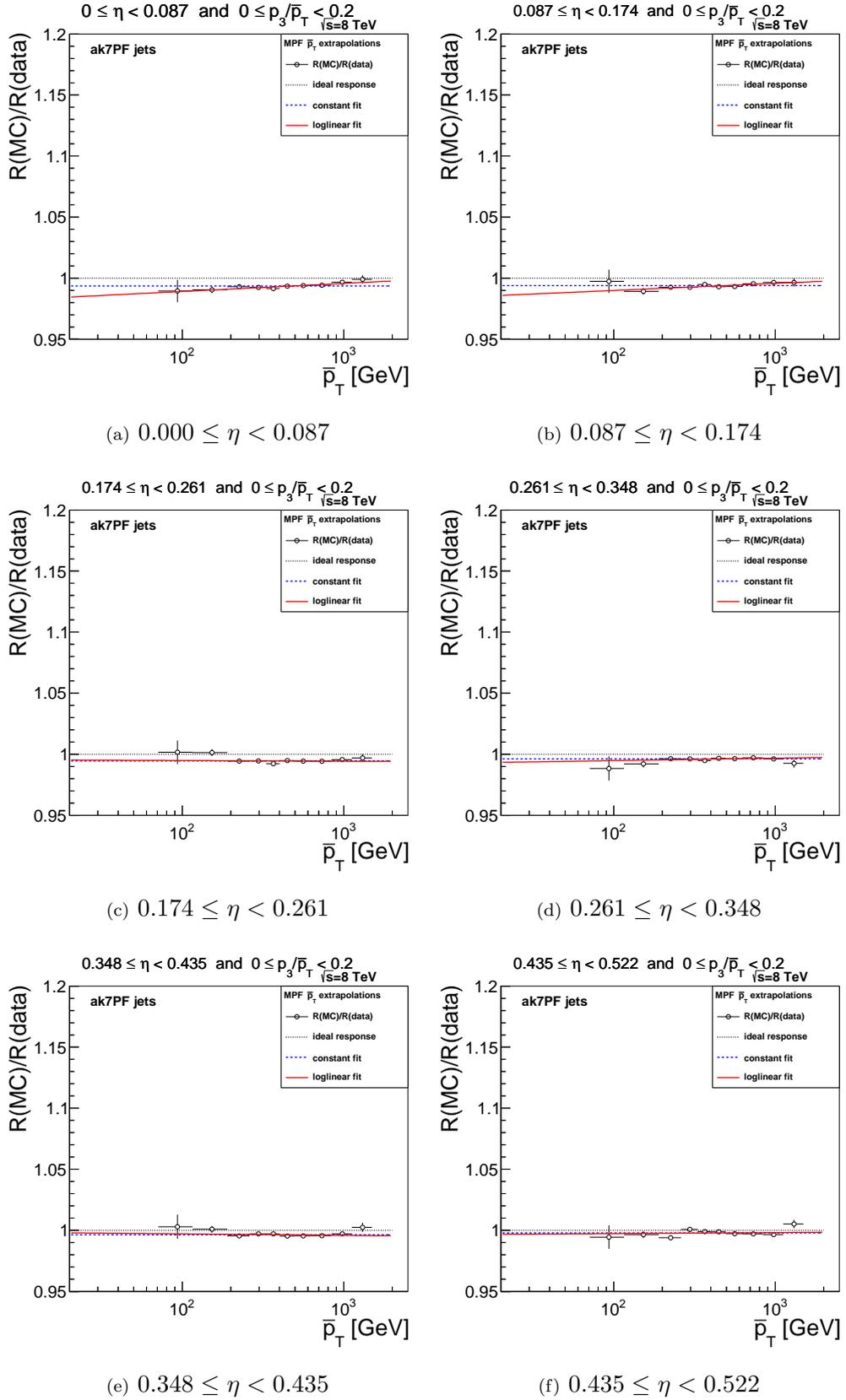
FIGURE A.41: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $0.000 \leq \eta < 0.522$ for ak7PF jets
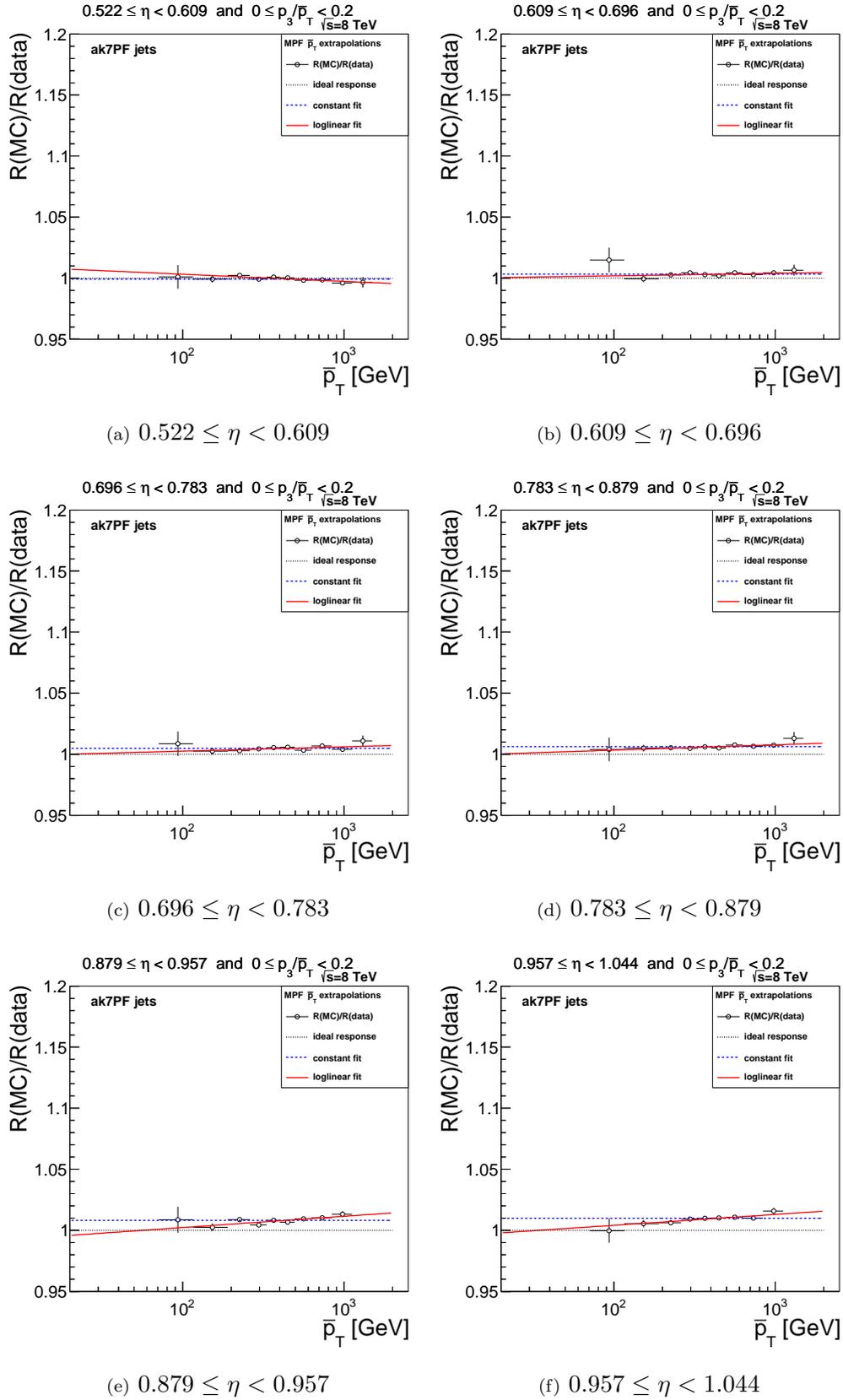
FIGURE A.42: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $0.522 \leq \eta < 1.044$ for ak7PF jets
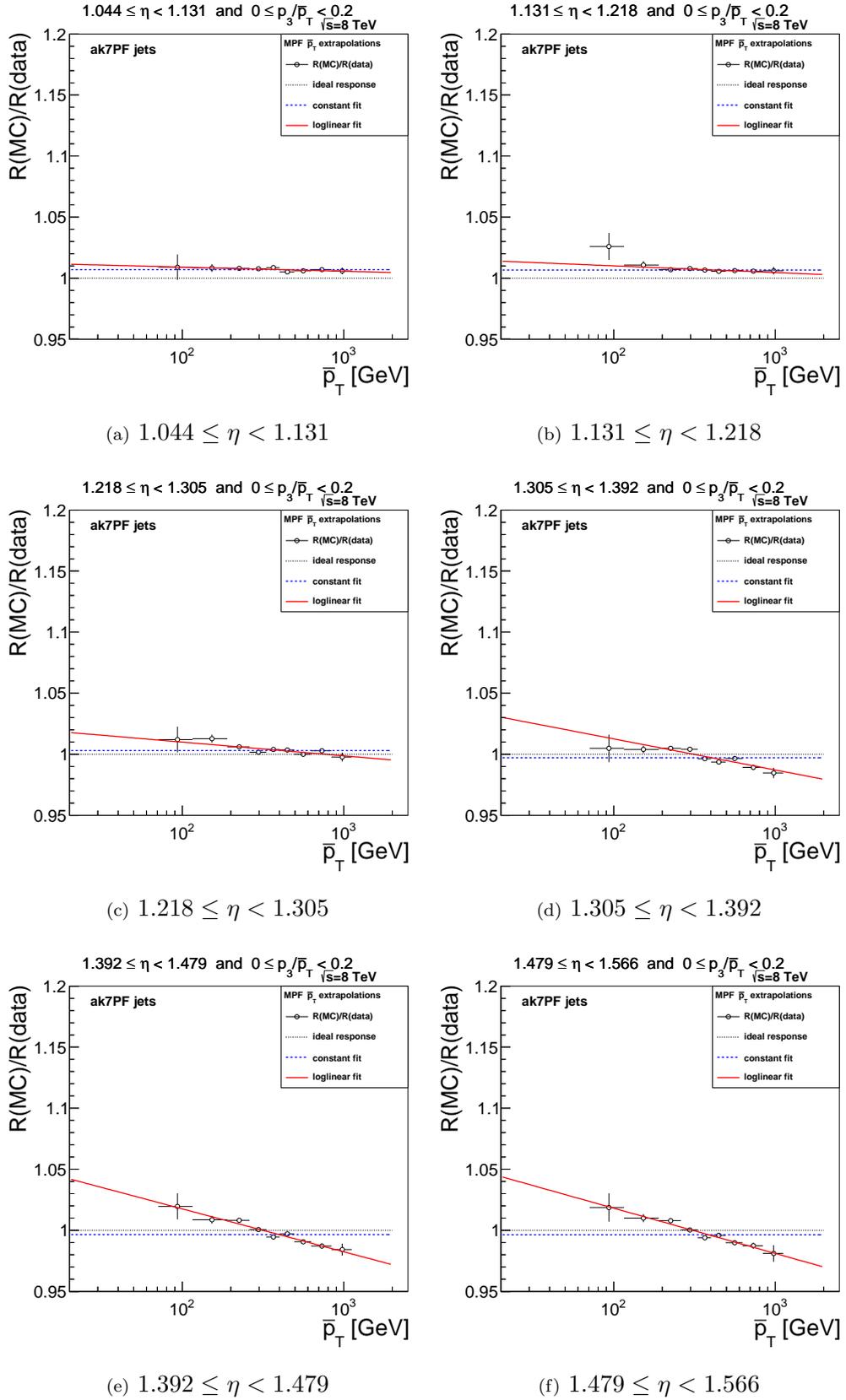
FIGURE A.43: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $1.044 \leq \eta < 1.566$ for ak7PF jets
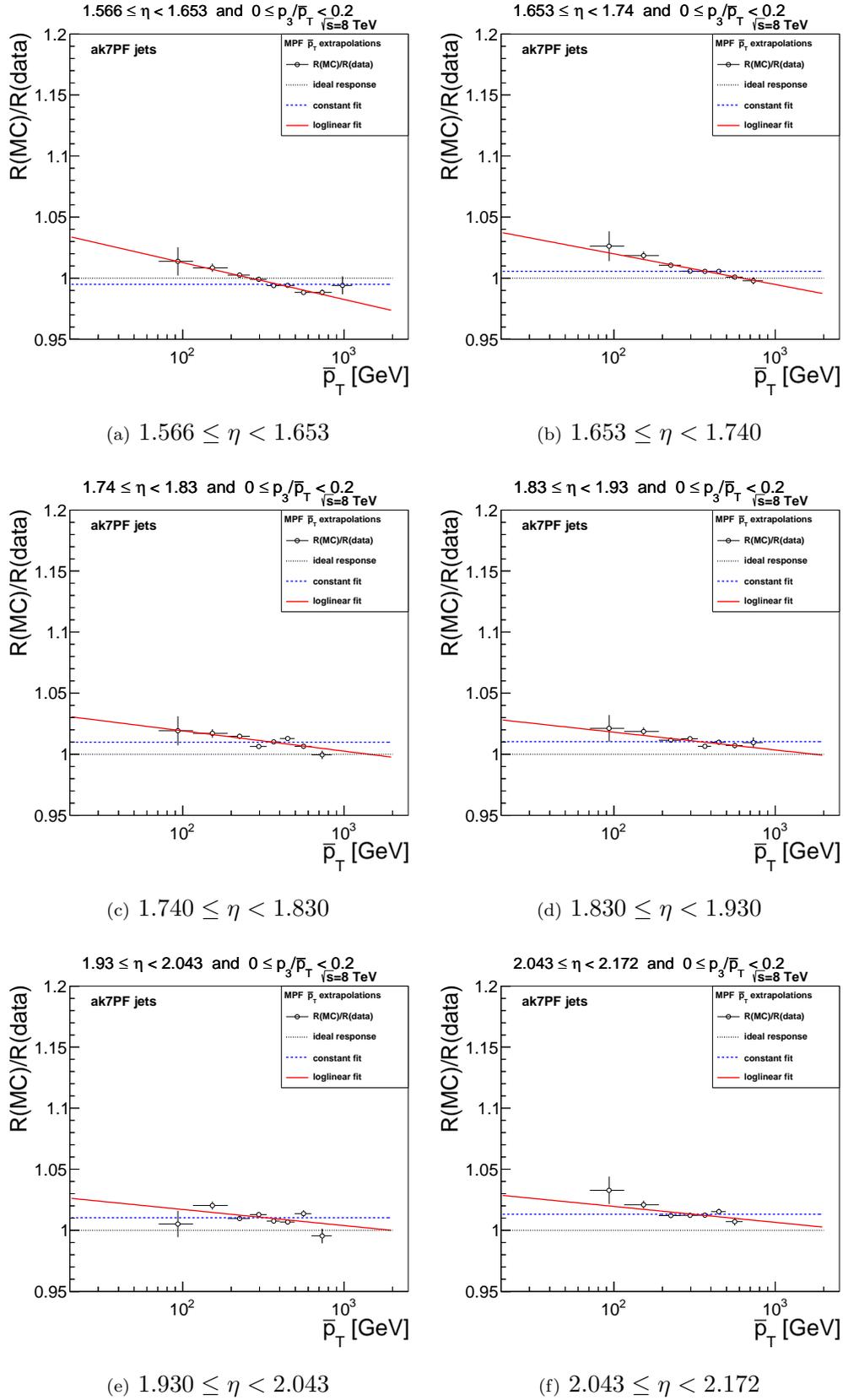
FIGURE A.44: Linear and loglinear extrapolations per η bin over $\bar{p}_T$ in $1.566 \leq \eta < 2.172$ for ak7PF jets

(a) $2.172 \leq \eta < 2.322$

(b) $2.322 \leq \eta < 2.500$

(c) $2.500 \leq \eta < 2.853$

(d) $2.853 \leq \eta < 2.964$

(e) $2.964 \leq \eta < 3.139$

(f) $3.139 \leq \eta < 5.191$
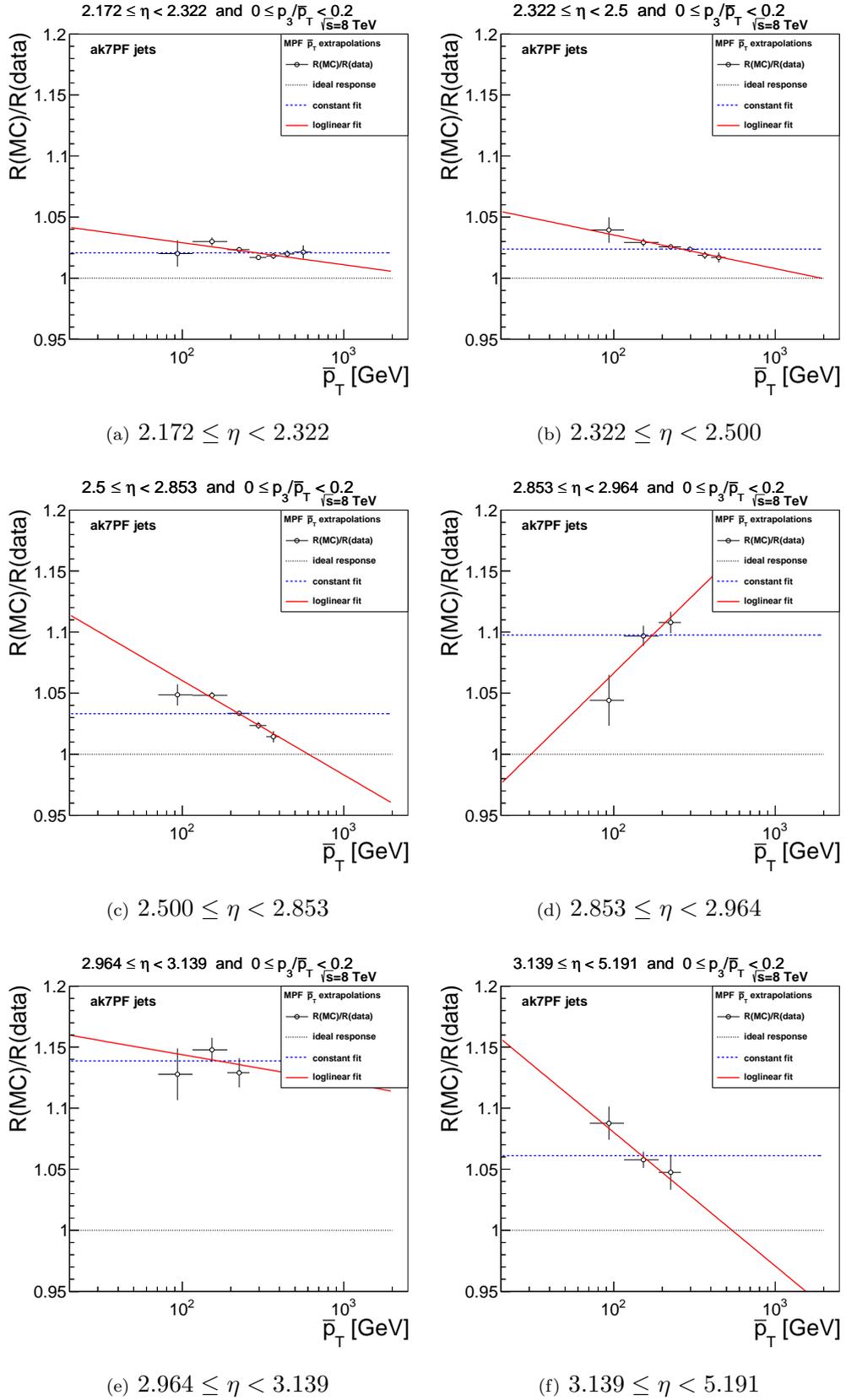
FIGURE A.45: Linear and loglinear extrapolations per η bin over $\overline{p}_T$ in $2.172 \leq \eta < 5.191$ for ak7PF jets

Appendix B

VBF search control plots

B.1 Efficiency Maps

(a) uds-quarks

(b) charm quarks

(c) bottom quarks

(d) gluons

(e) unmatched
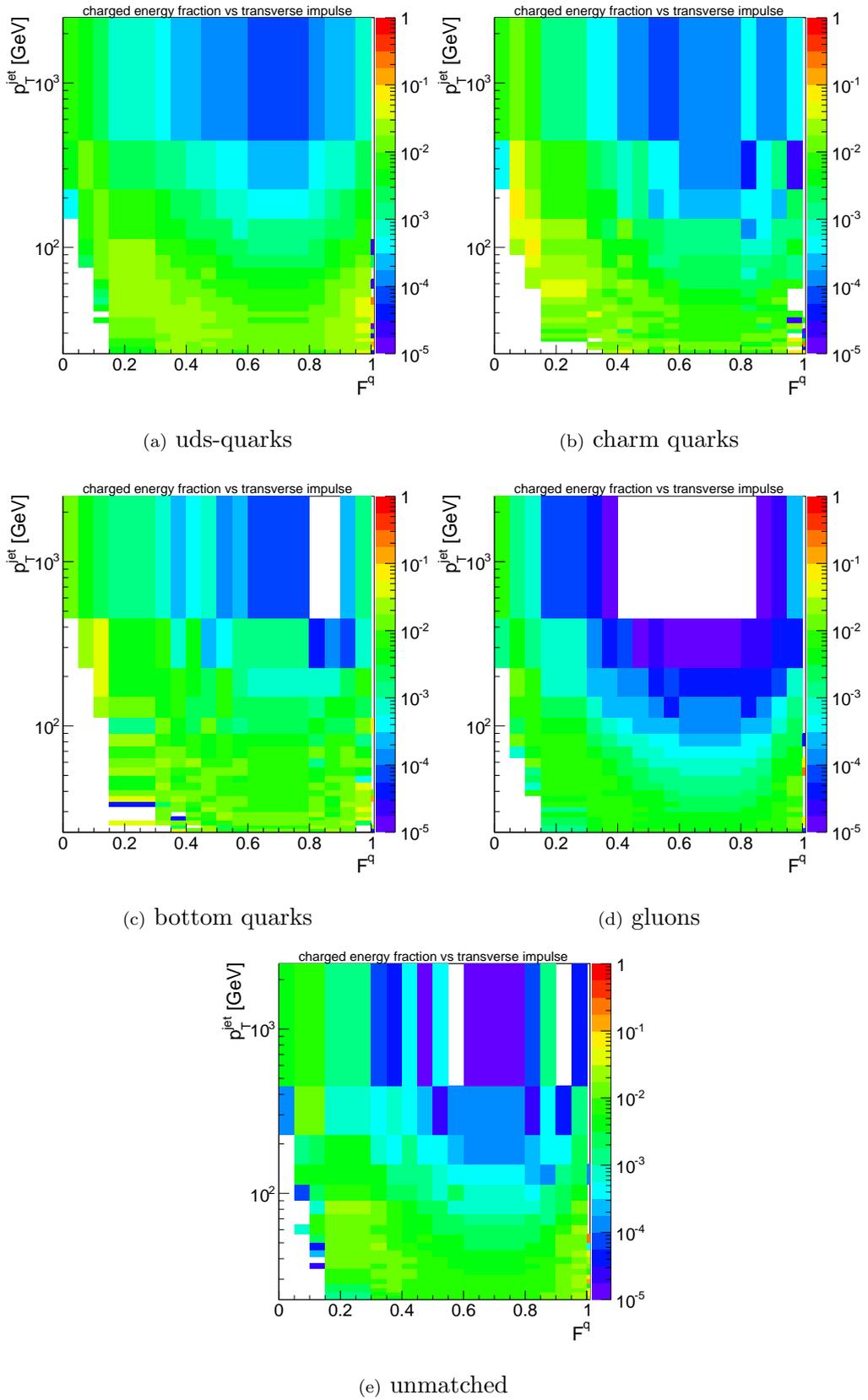
FIGURE B.1: Maps for tight isolation fakes

(a) uds-quarks

(b) charm quarks

(c) bottom quarks

(d) gluons

(e) unmatched
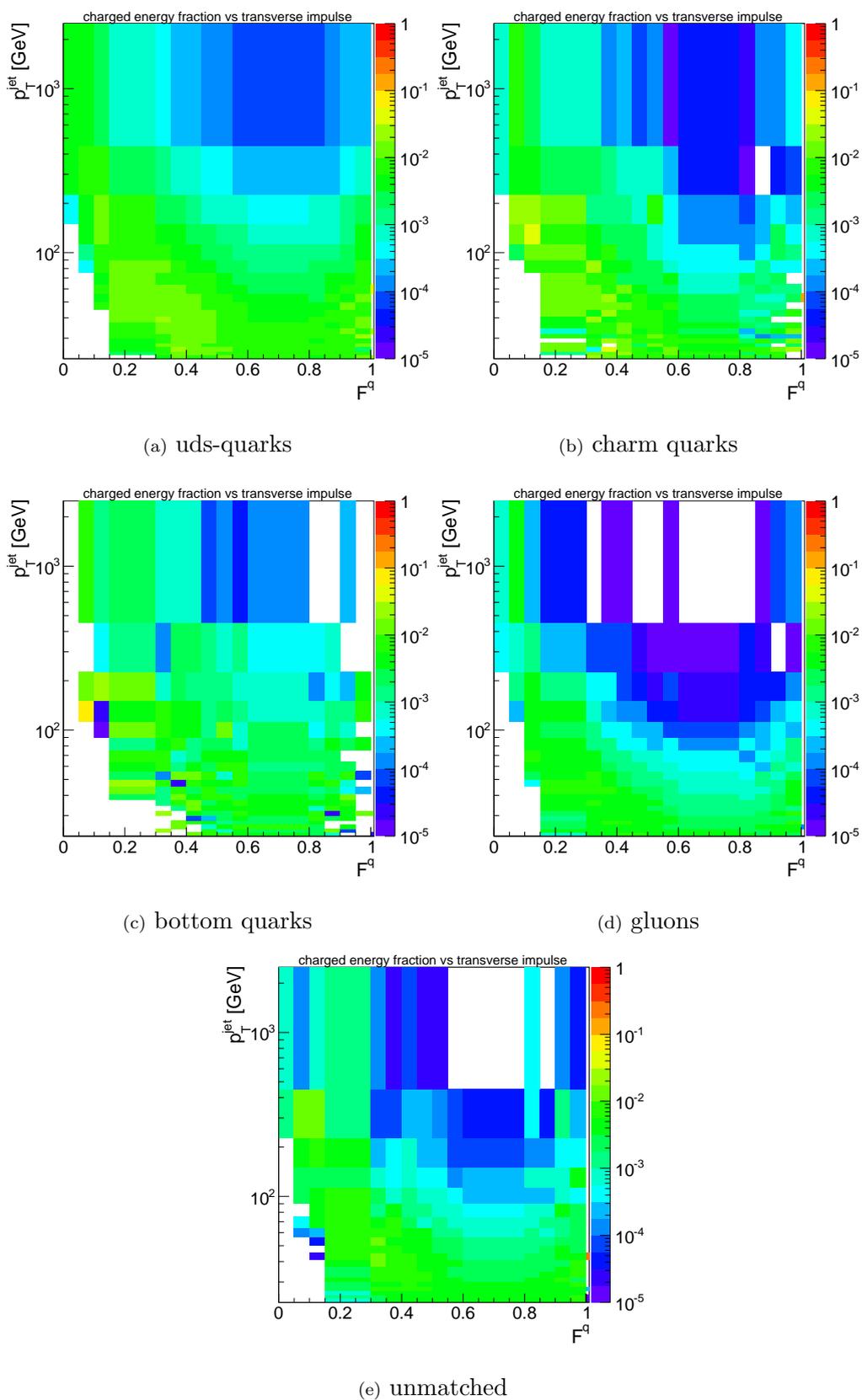
FIGURE B.2: Maps for exclusive medium isolation fakes

(a) uds-quarks

(b) charm quarks

(c) bottom quarks

(d) gluons

(e) unmatched

FIGURE B.3: Maps for exclusive loose isolation fakes

B.2   Selection acceptance correction

(a) tight

(b) medium

(c) loose

(d) inclusive
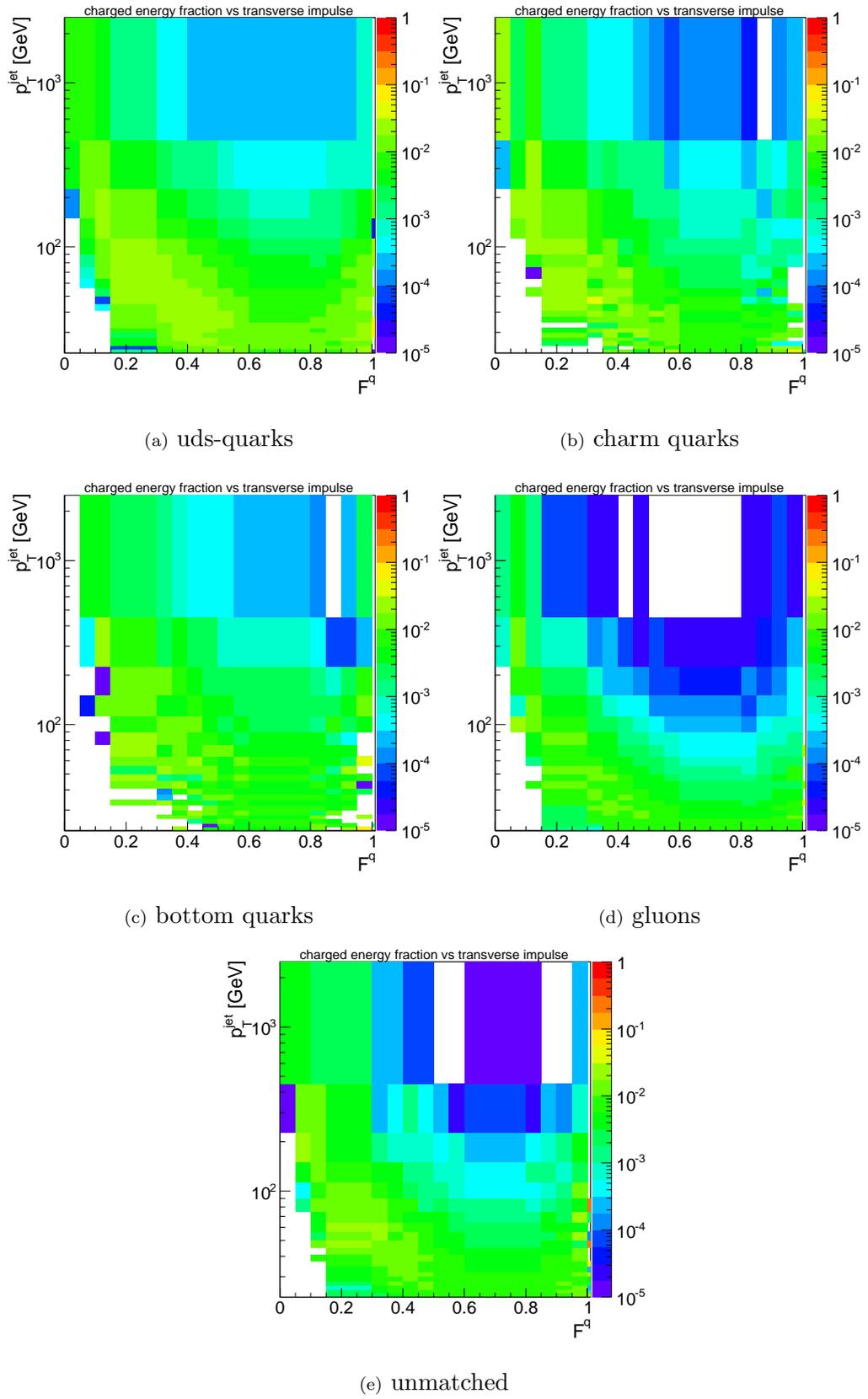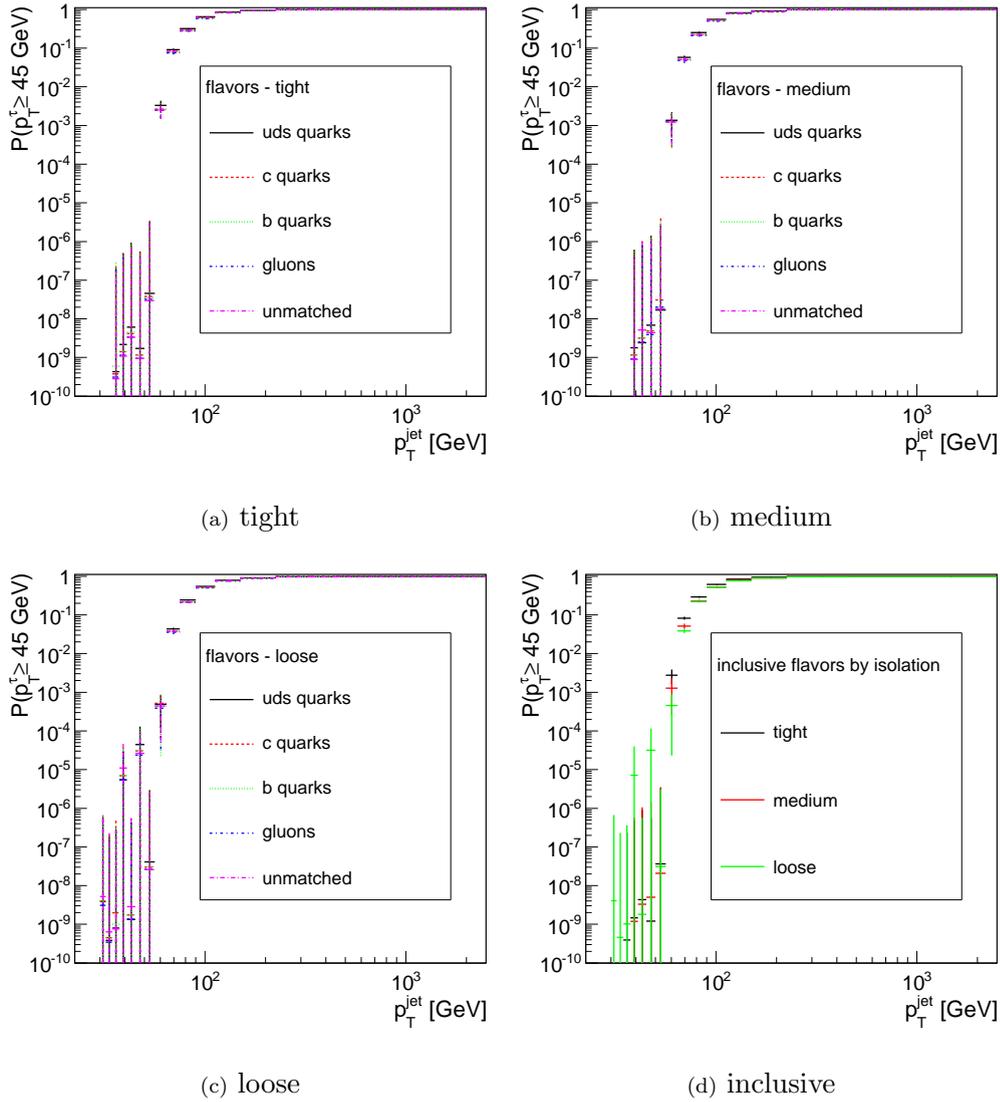
FIGURE B.4: Probability of a jet with a given p_T^{jet} to fake a τ passing selection thresholds for a given τ^{fake}-isolation for different flavours

B.3 Jet to τ_h transverse momentum scale factors

(a) tight

(b) medium

(c) loose

(d) inclusive
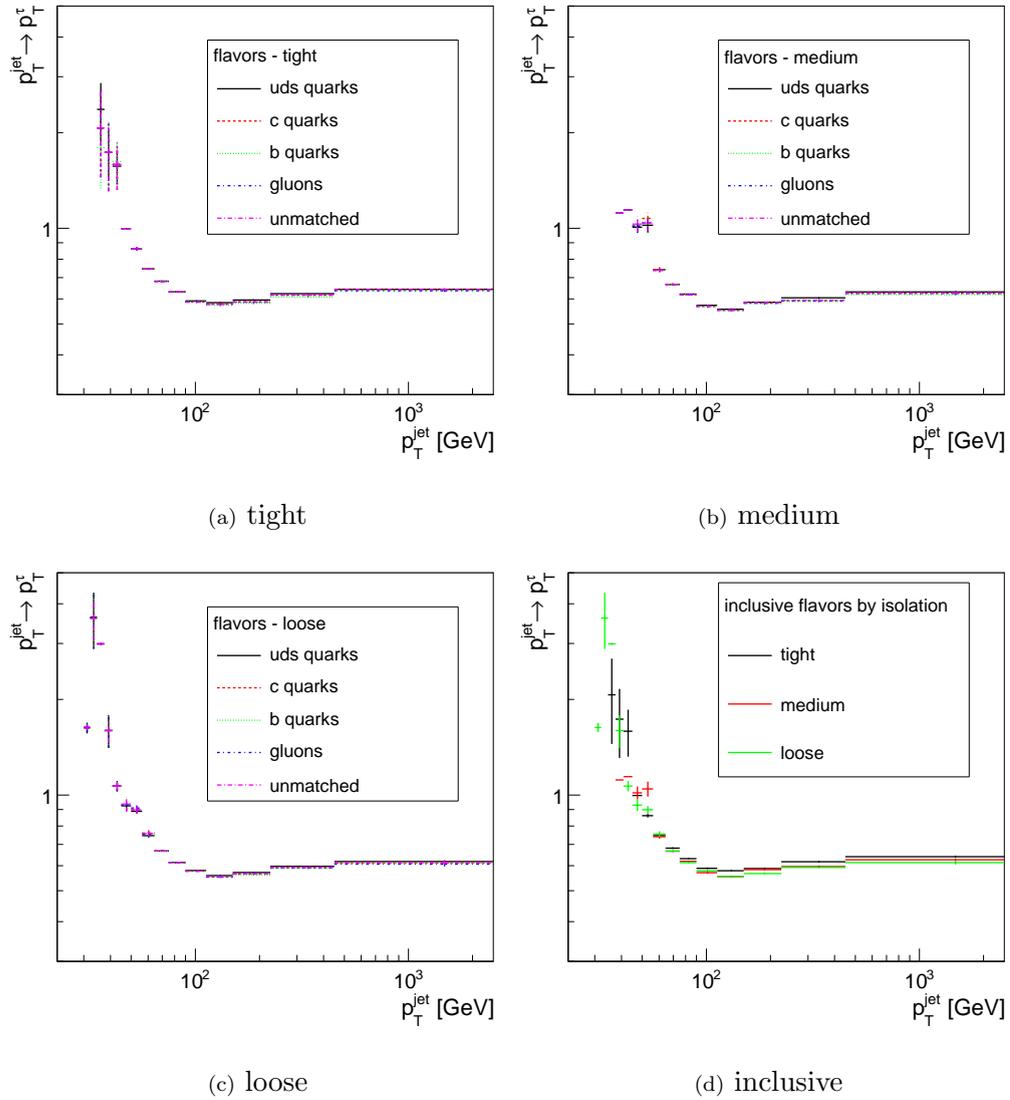
FIGURE B.5: Average translation factor of a jet with a given p_T^{jet} faking a τ of a given isolation to yield $p_T^{\tau^{fake}}$ after reconstruction as a τ^{fake}

B.4 Data to MC comparisons

B.4.1 Cutflows

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
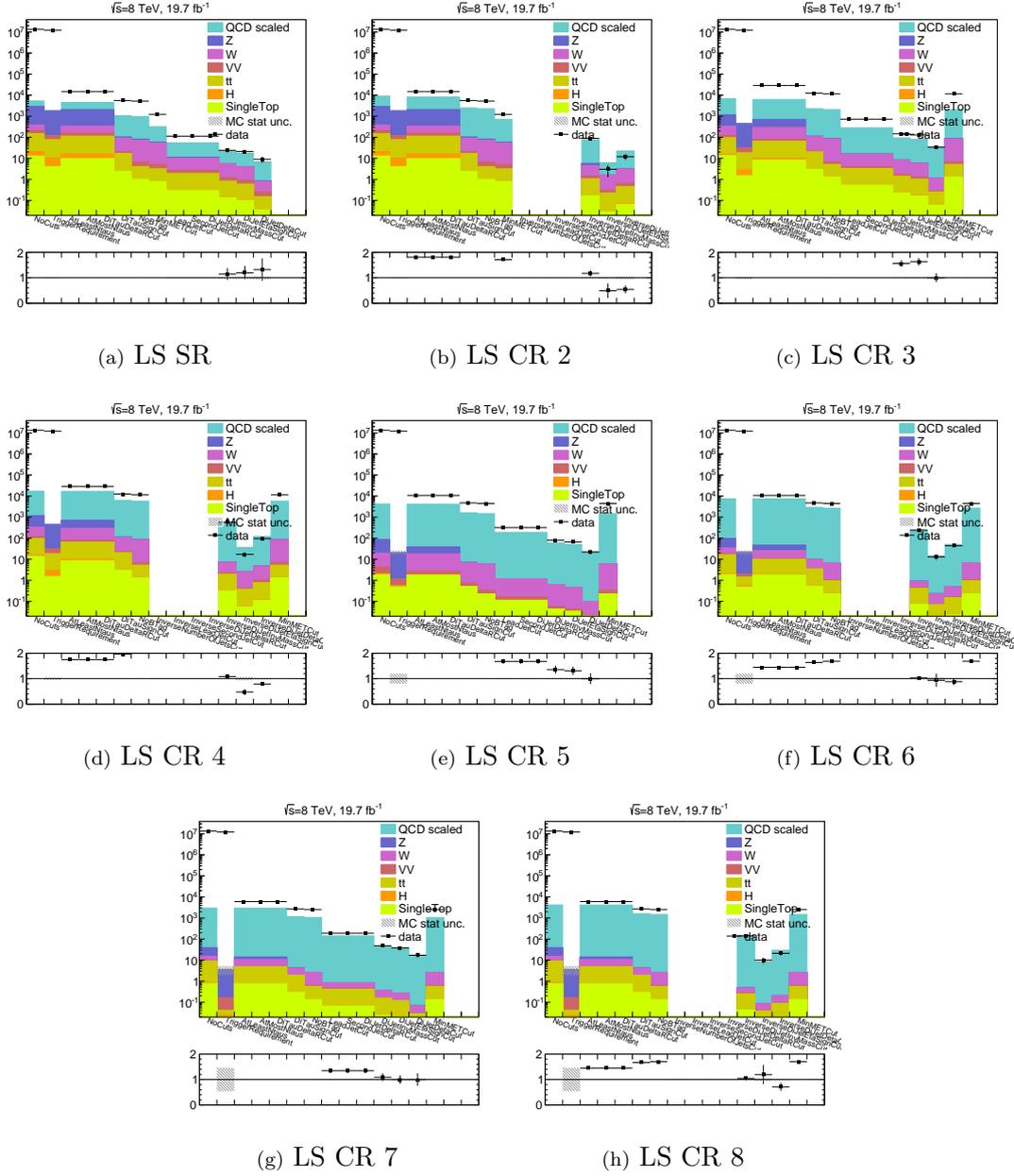
FIGURE B.6: Logarithmic cutflows for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
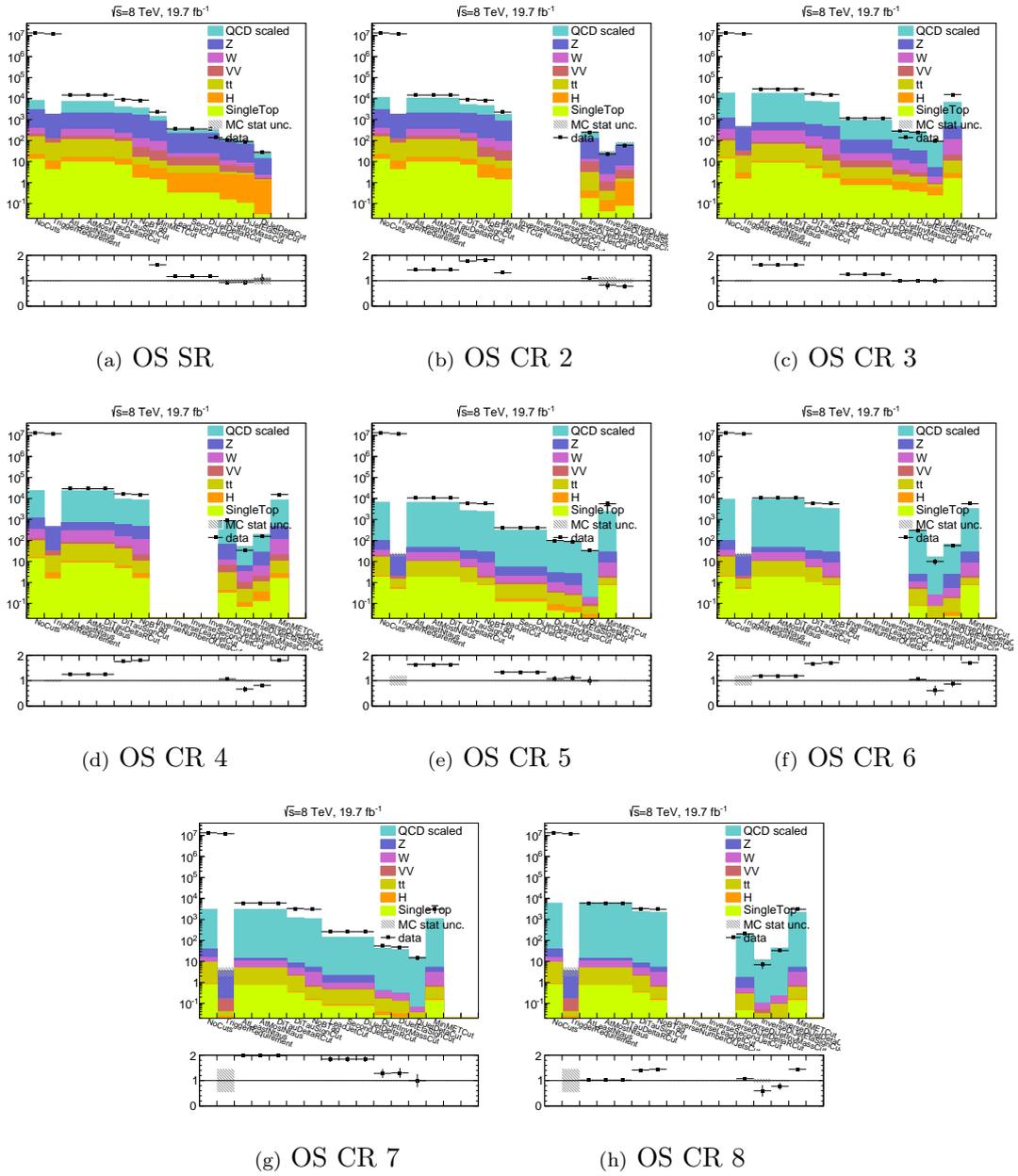
FIGURE B.7: Logarithmic cutflows for OS

B.4.2 $\Delta\eta(jet, jet)$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
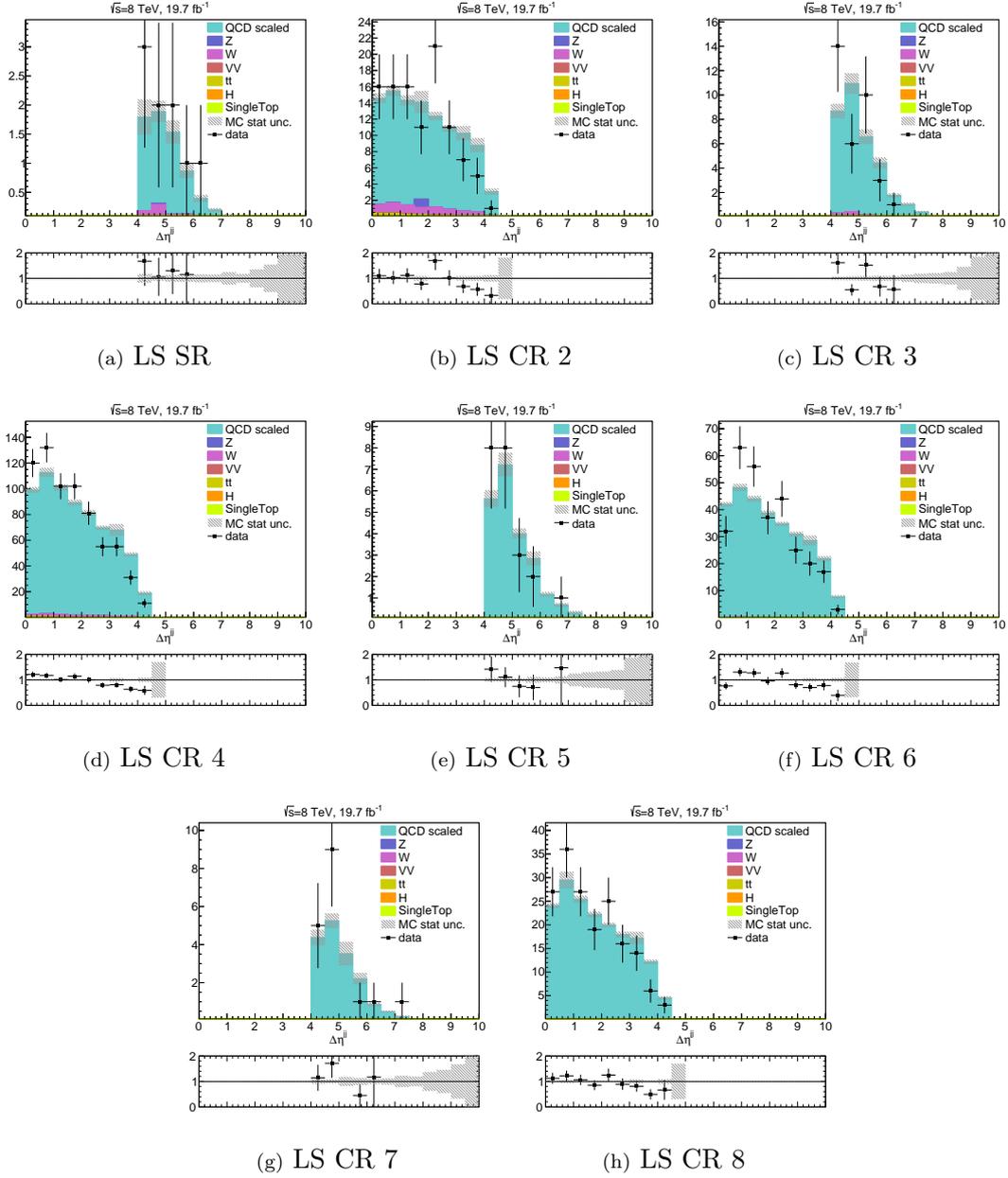
FIGURE B.8: Linear difference in η for the vector boson fusion dijet system for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
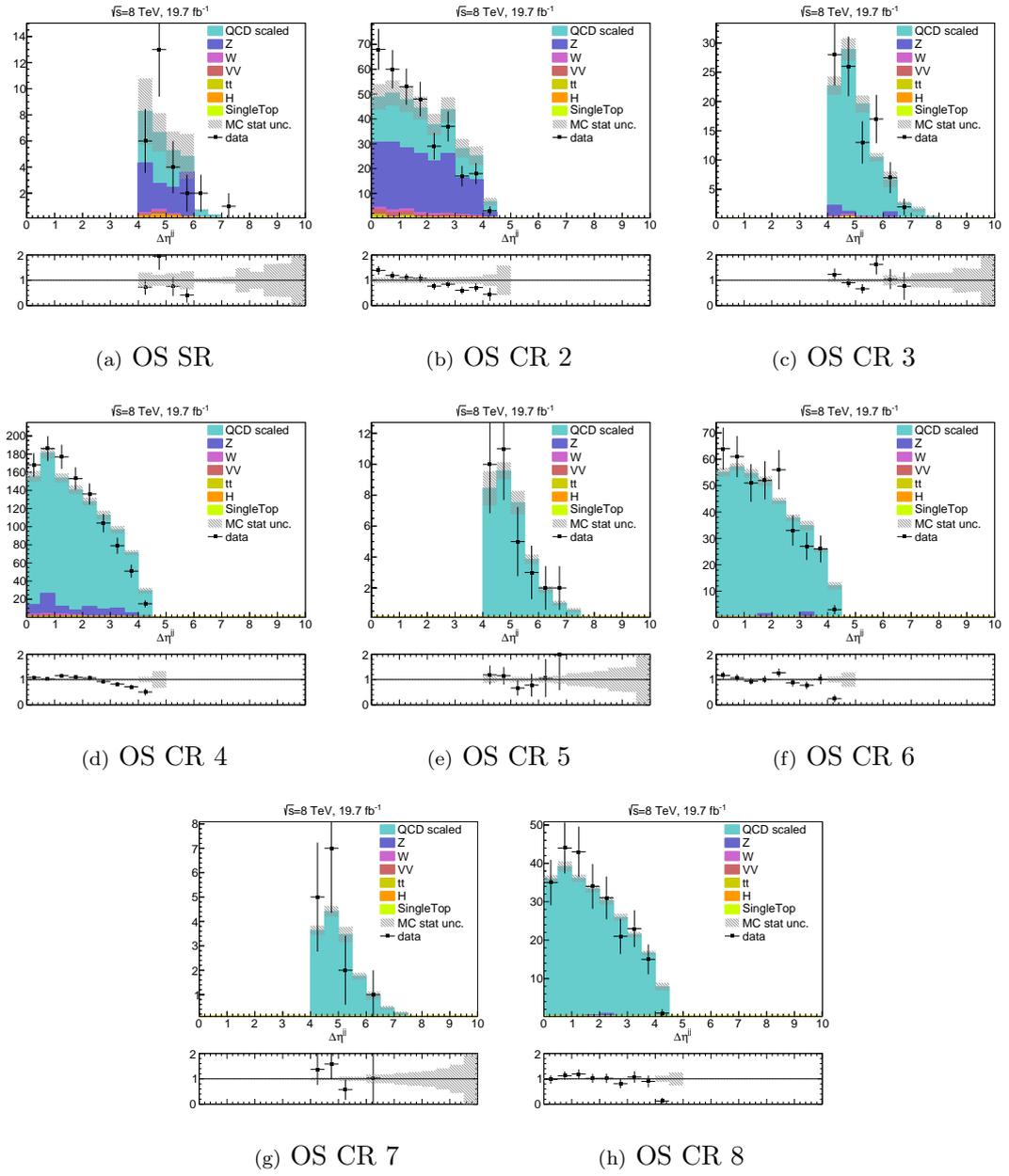
FIGURE B.9: Linear difference in η for the vector boson fusion dijet system for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
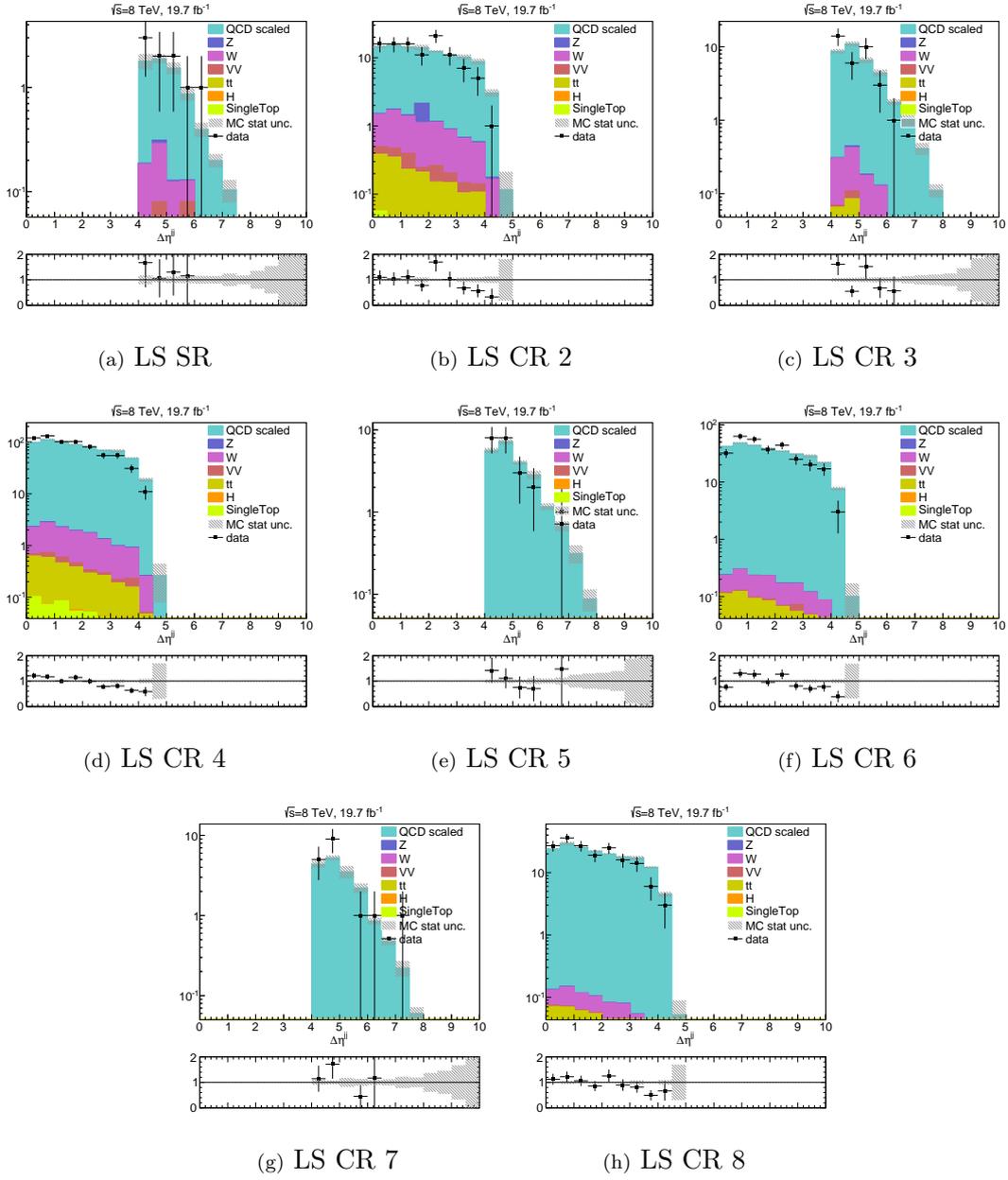
FIGURE B.10: Logarithmic difference in η for the vector boson fusion dijet system for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
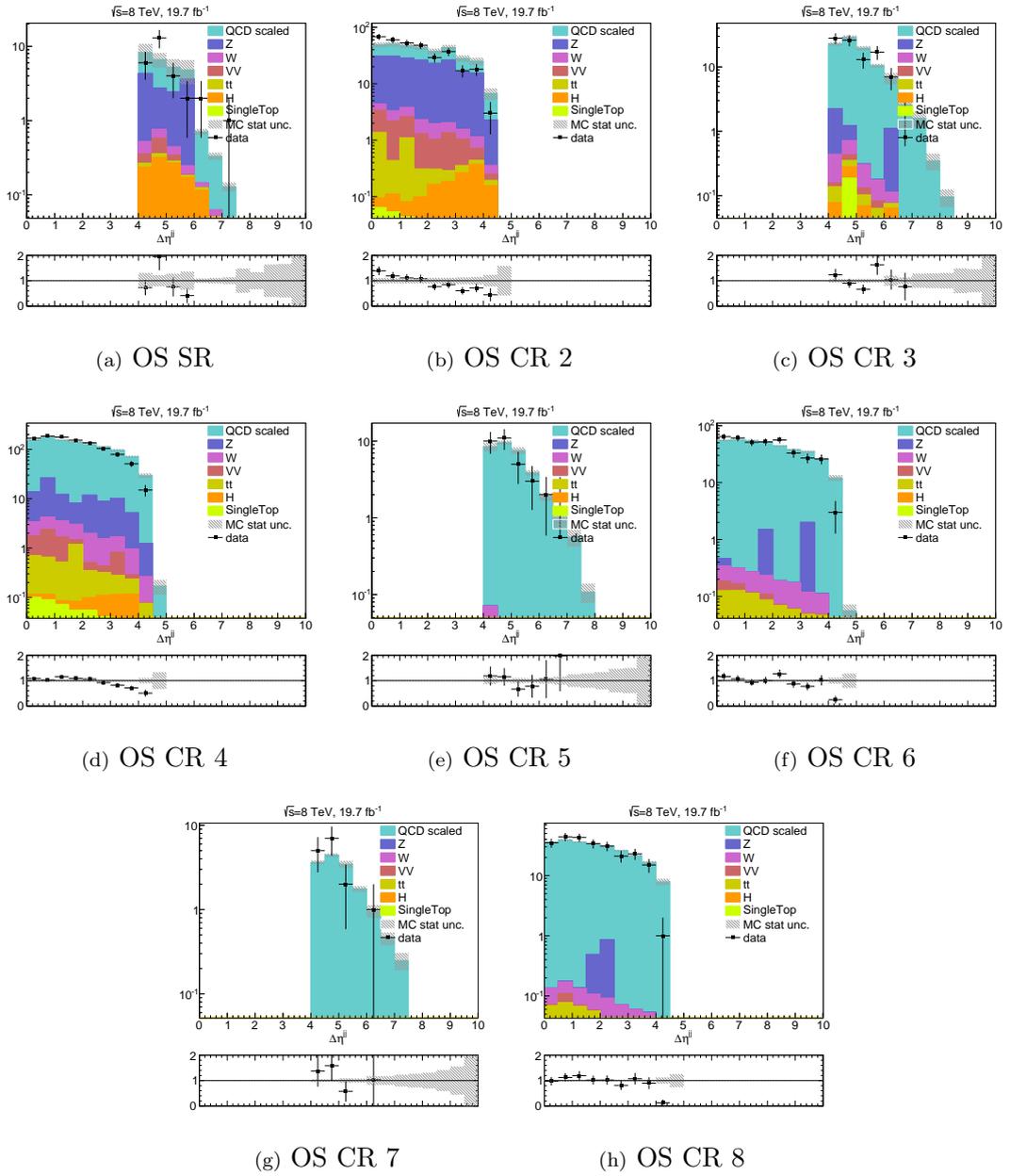
FIGURE B.11: Logarithmic difference in η for the vector boson fusion dijet system for OS

B.4.3 $M^{jet,jet}$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
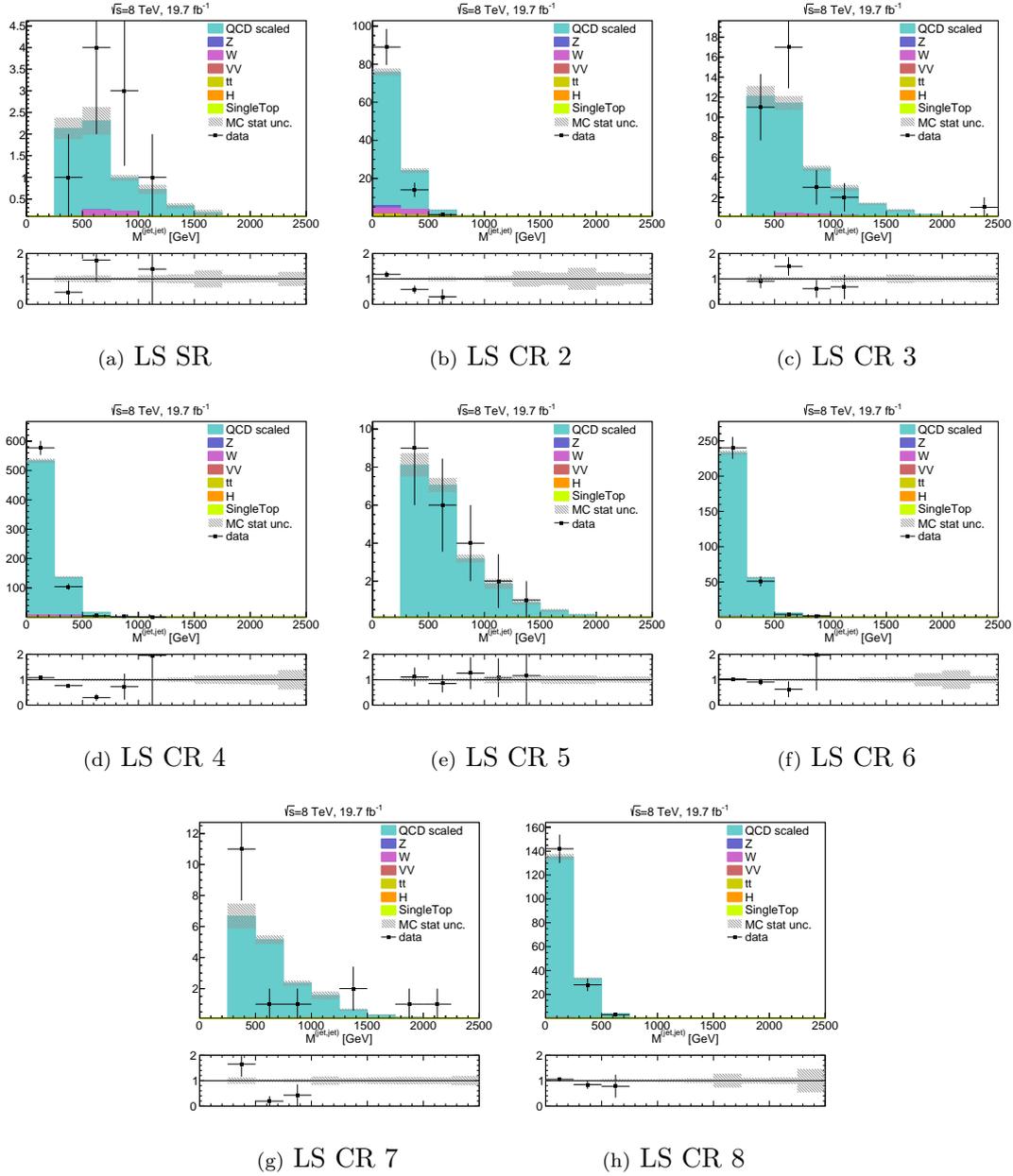
FIGURE B.12: Linear invariant mass of the vector boson fusion dijet system for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
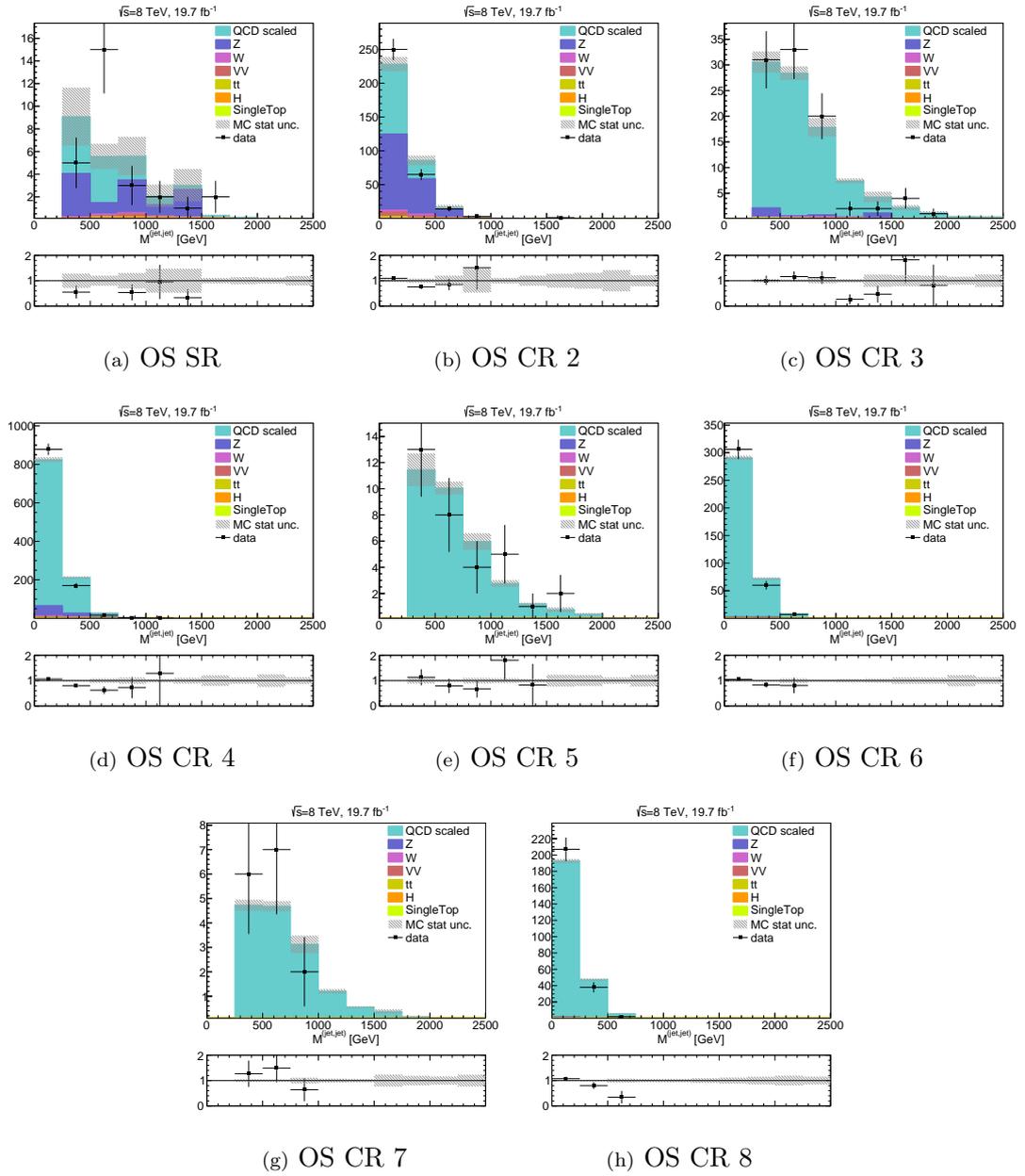
FIGURE B.13: Linear invariant mass of the vector boson fusion dijet system for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
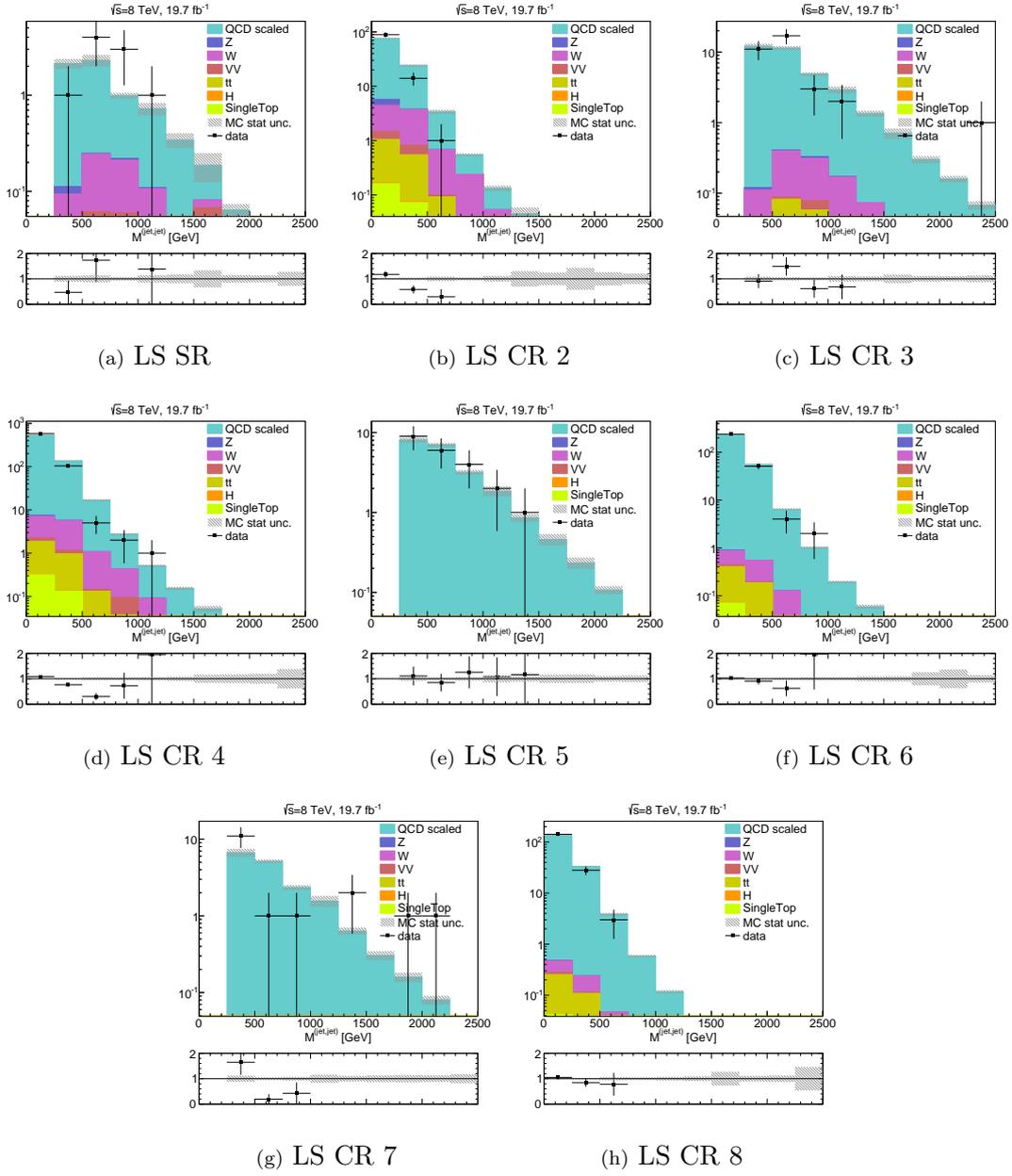
FIGURE B.14: Logarithmic invariant mass of the vector boson fusion dijet system for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
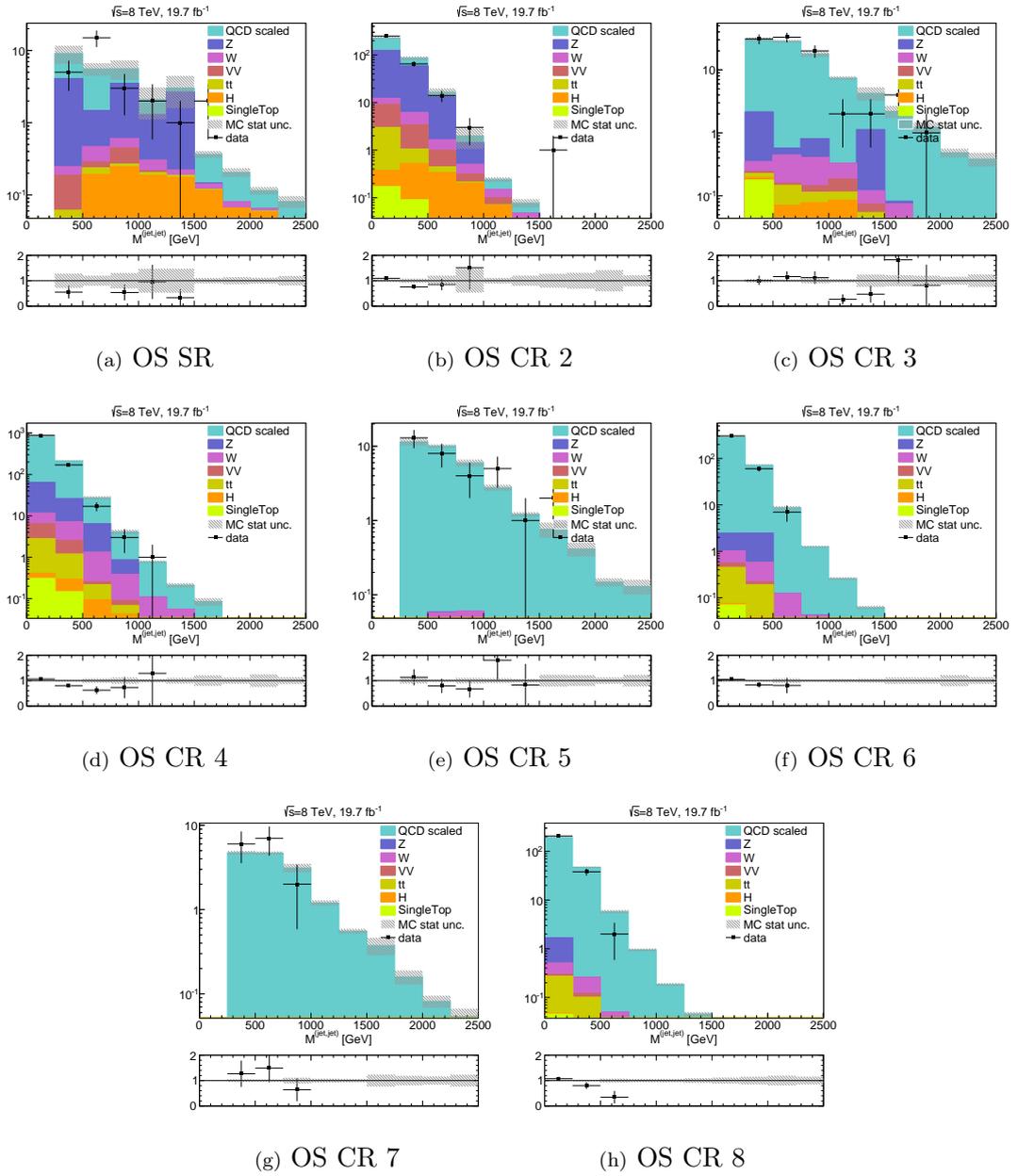
FIGURE B.15: Logarithmic invariant mass of the vector boson fusion dijet system for OS

B.4.4 $sign(q^{\tau\ 1} \cdot q^{\tau\ 2})$

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
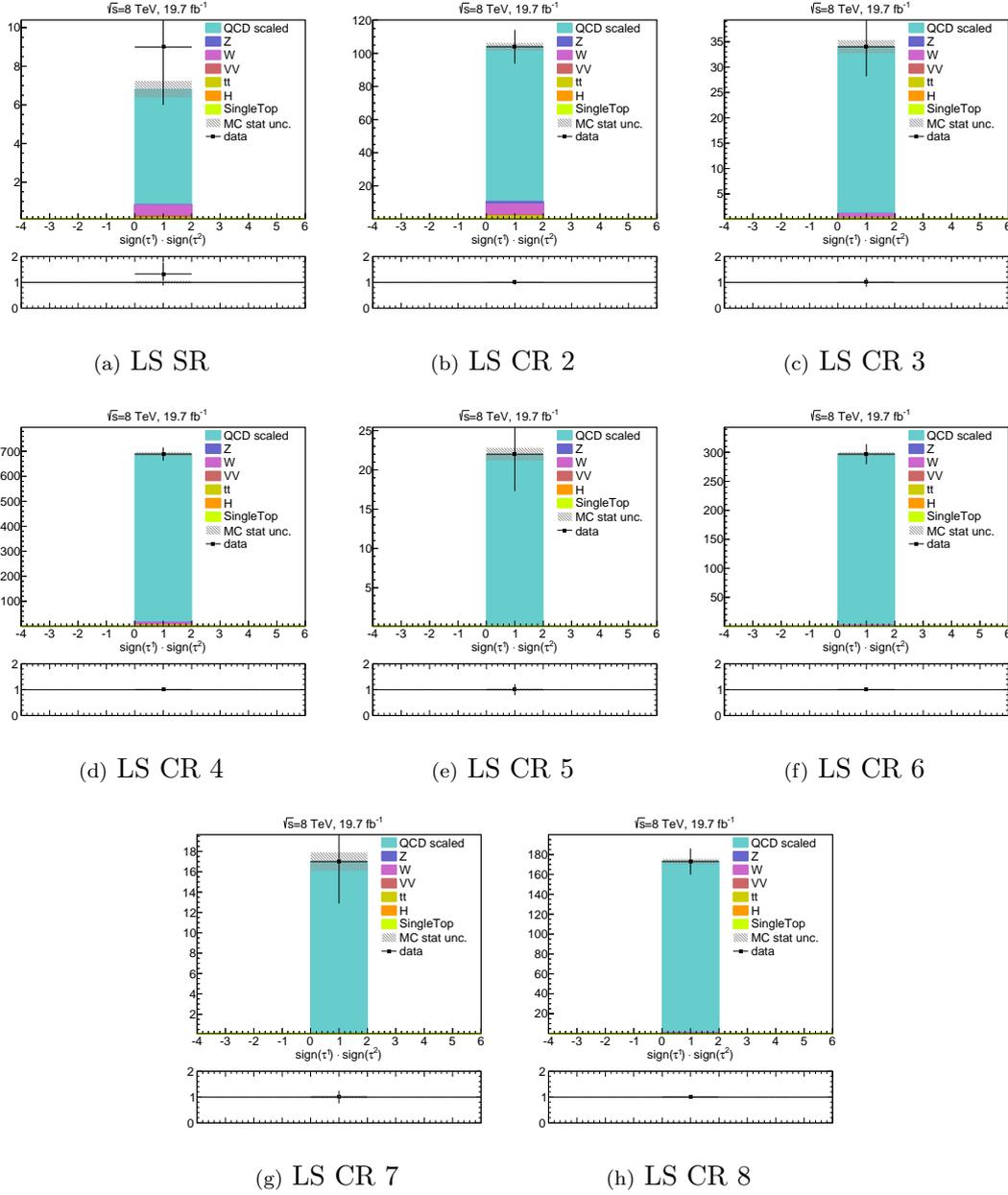
FIGURE B.16: Linear sign of the product of τ-charges for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
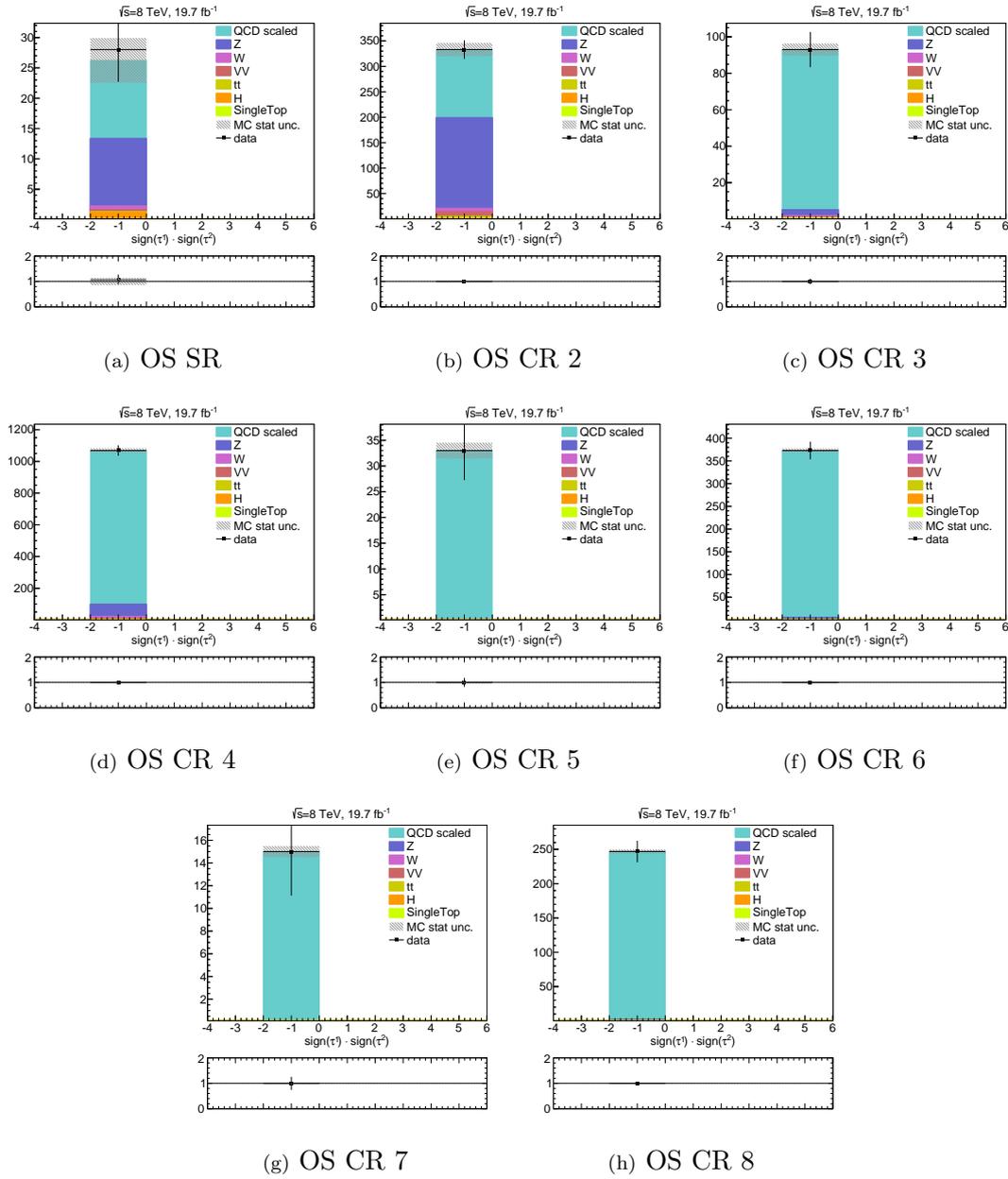
FIGURE B.17: Linear sign of the product of τ-charges for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
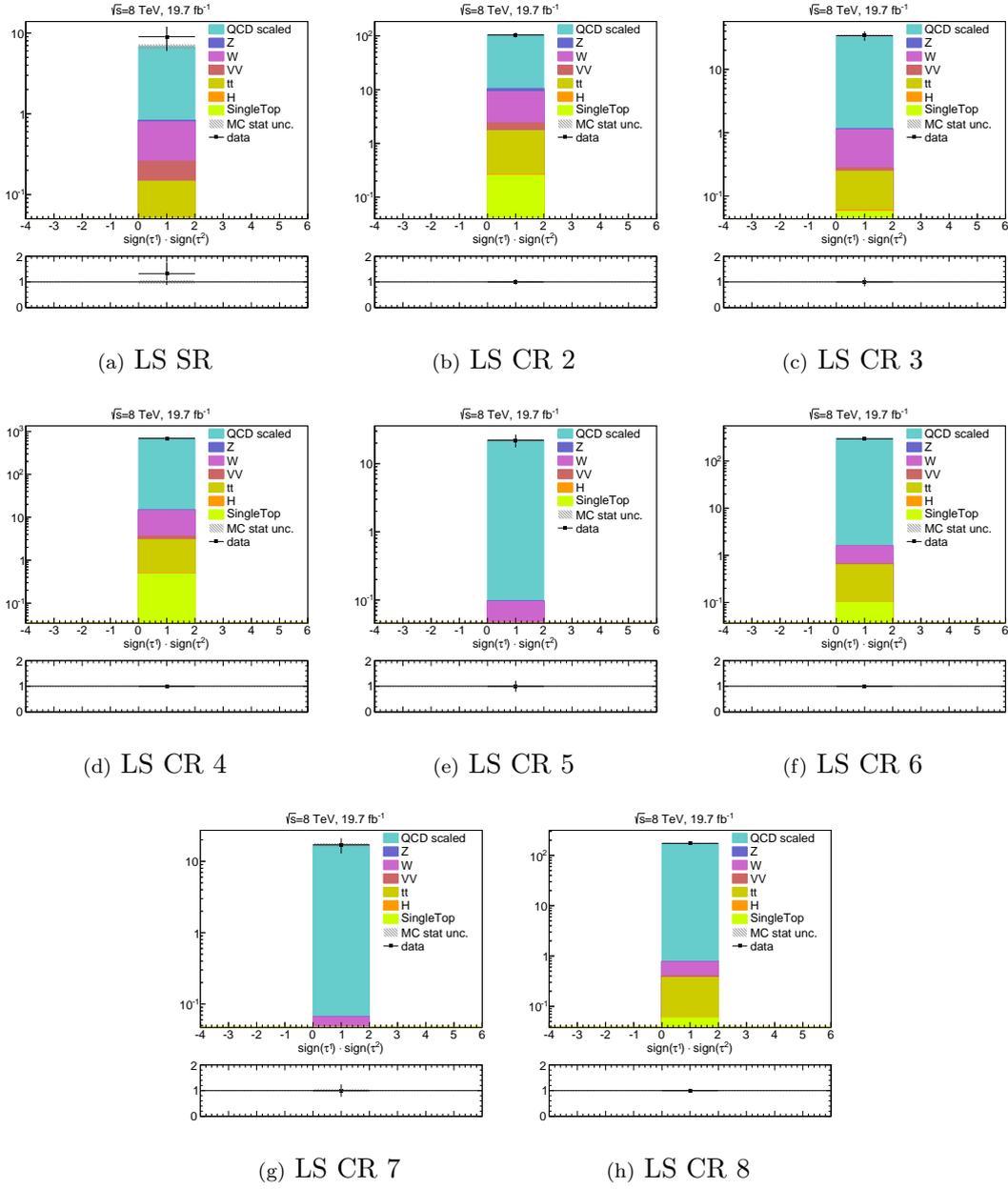
FIGURE B.18: Logarithmic sign of the product of τ-charges for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
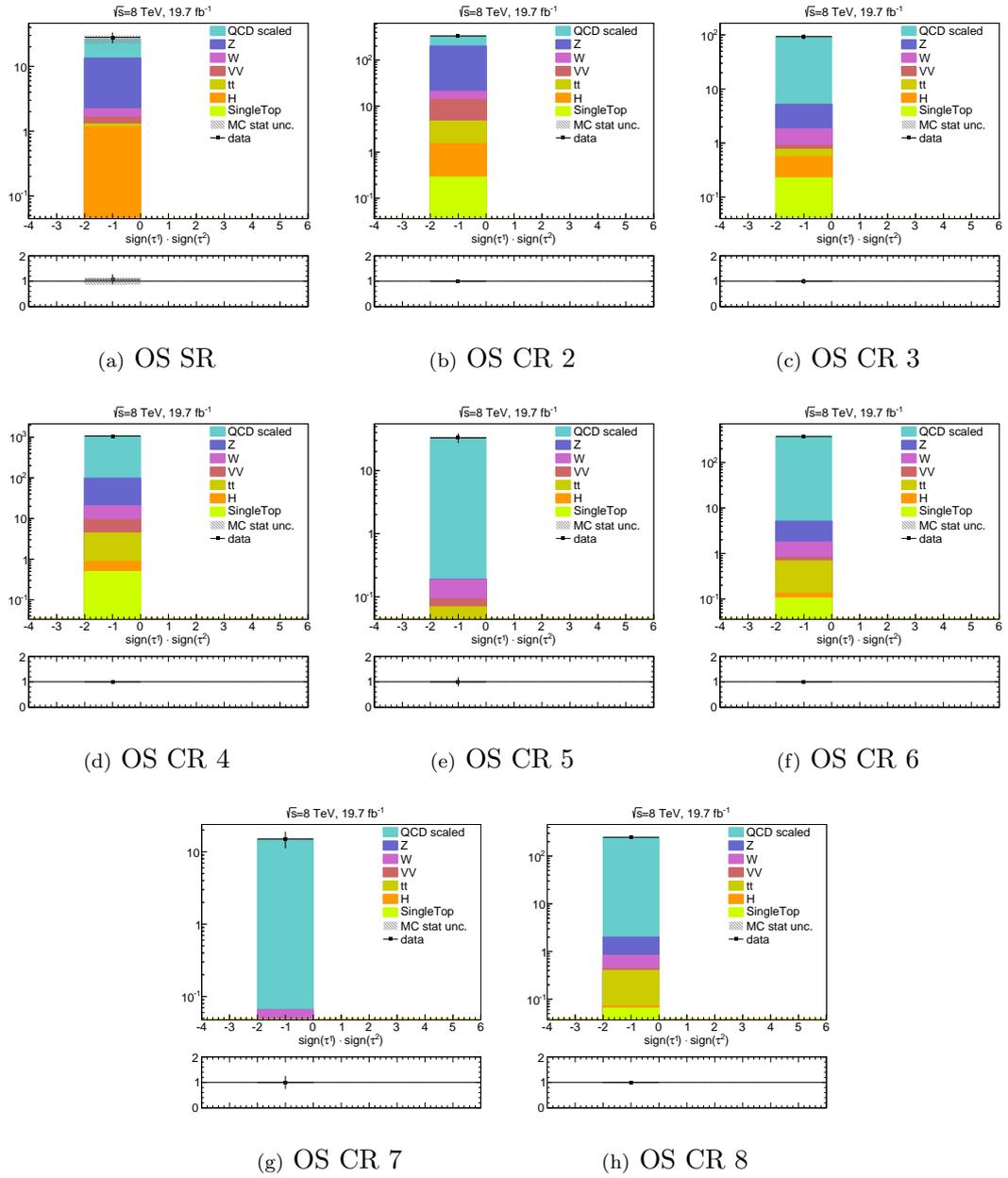
FIGURE B.19: Logarithmic sign of the product of τ-charges for OS

B.4.5 $\cos \Delta \phi(\tau, \tau)$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
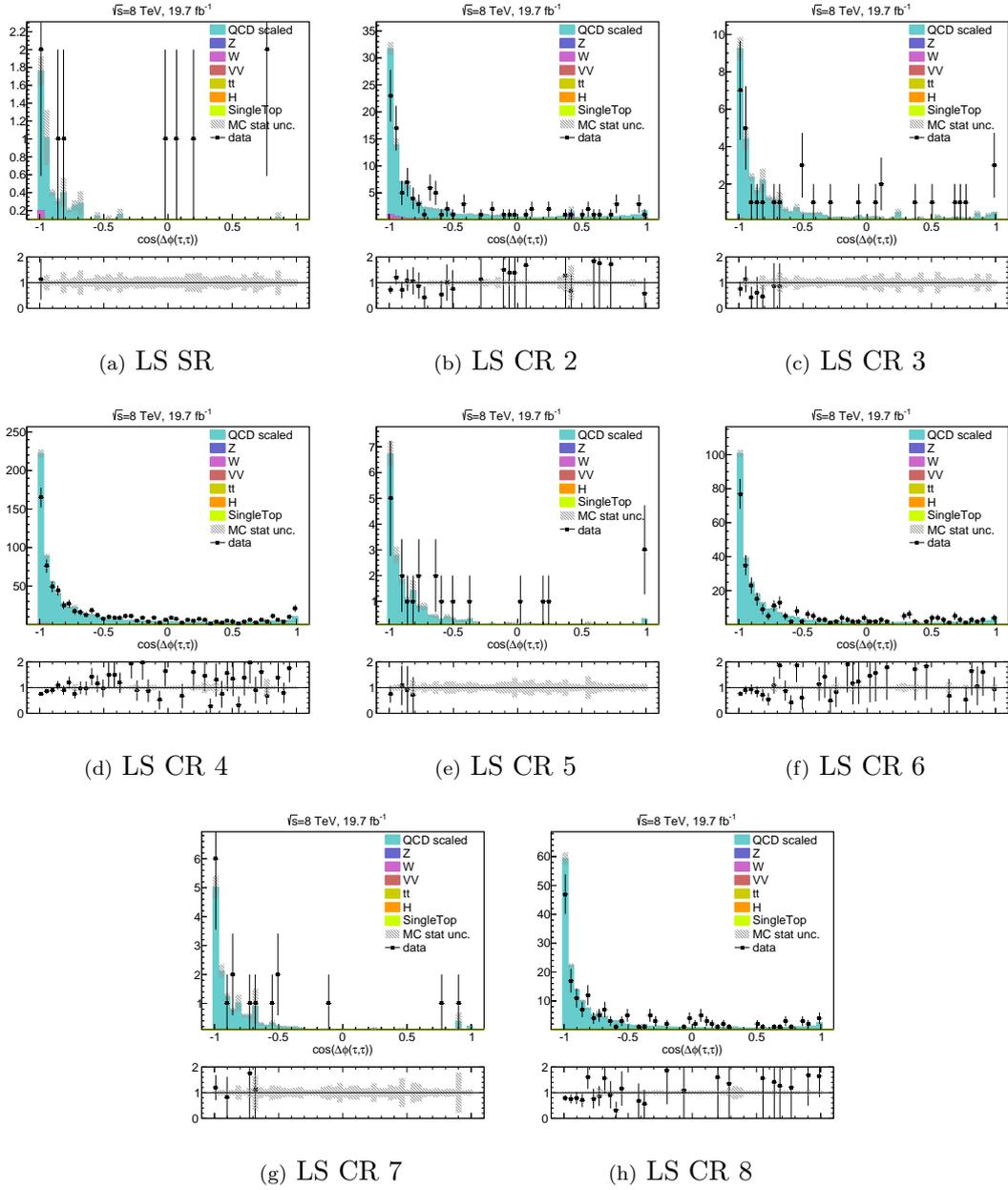
FIGURE B.20: Linear cosine of the difference of the τ in ϕ for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
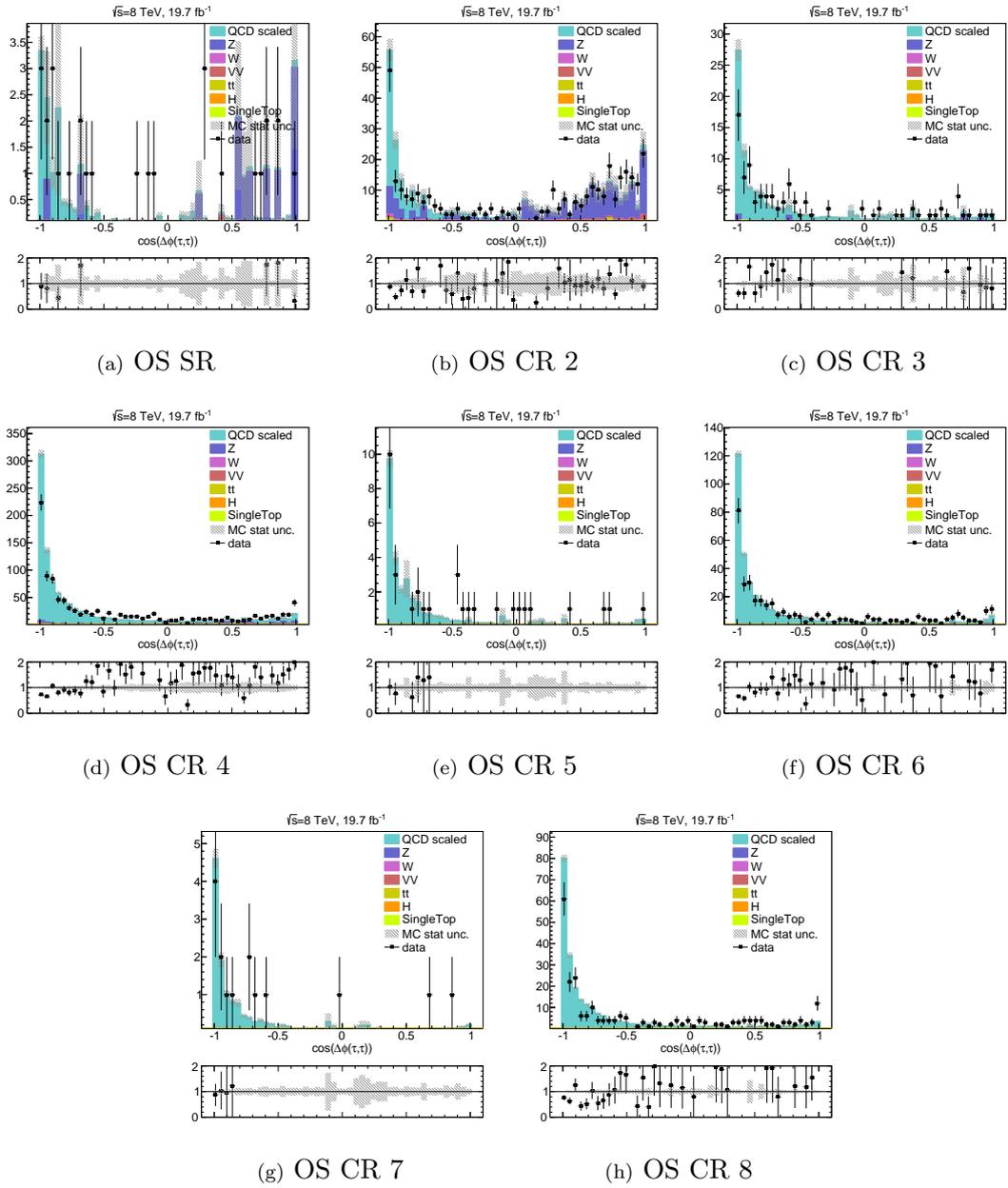
FIGURE B.21: Linear cosine of the difference of the τ in ϕ for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
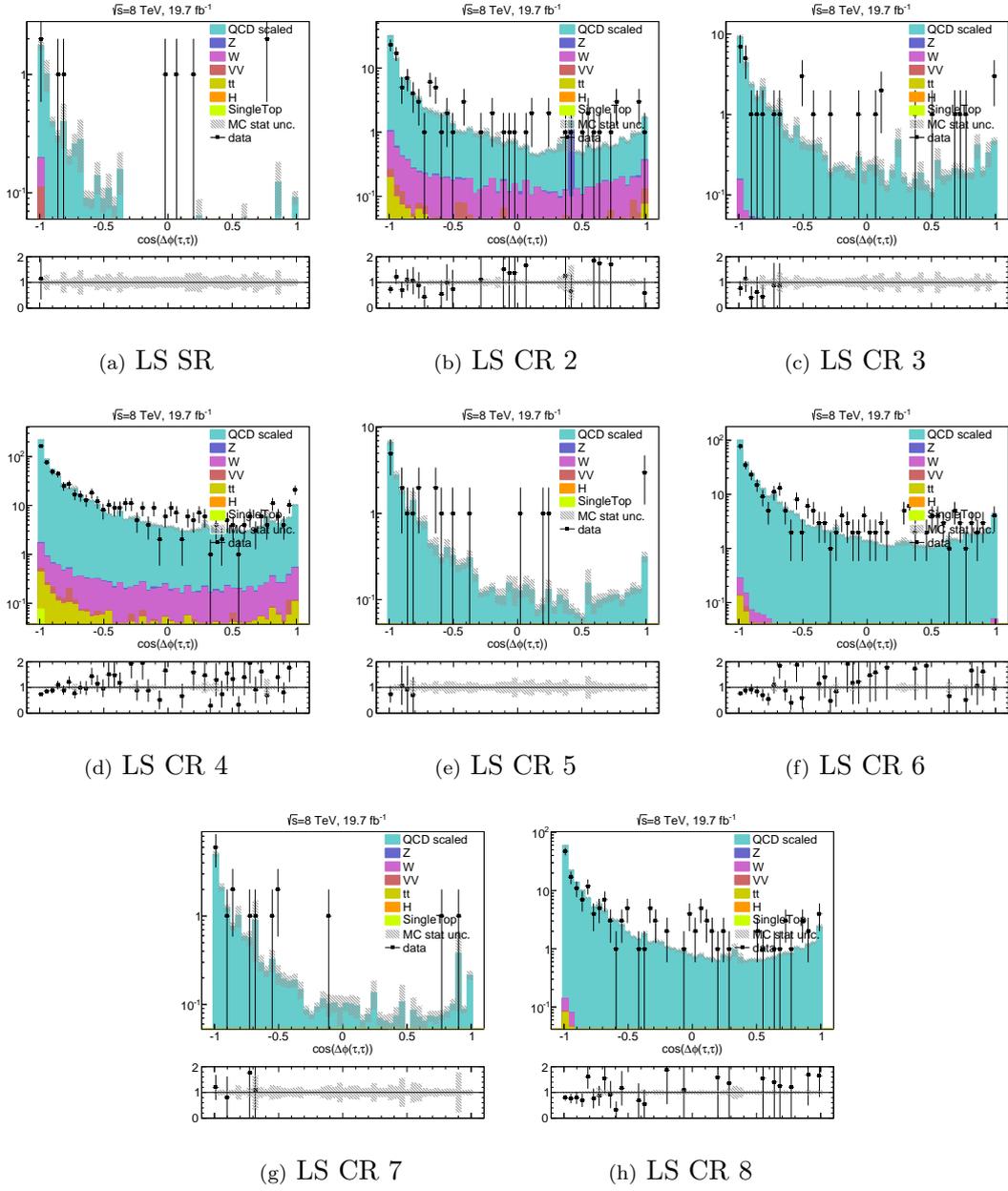
FIGURE B.22: Logarithmic cosine of the difference of the τ in ϕ for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
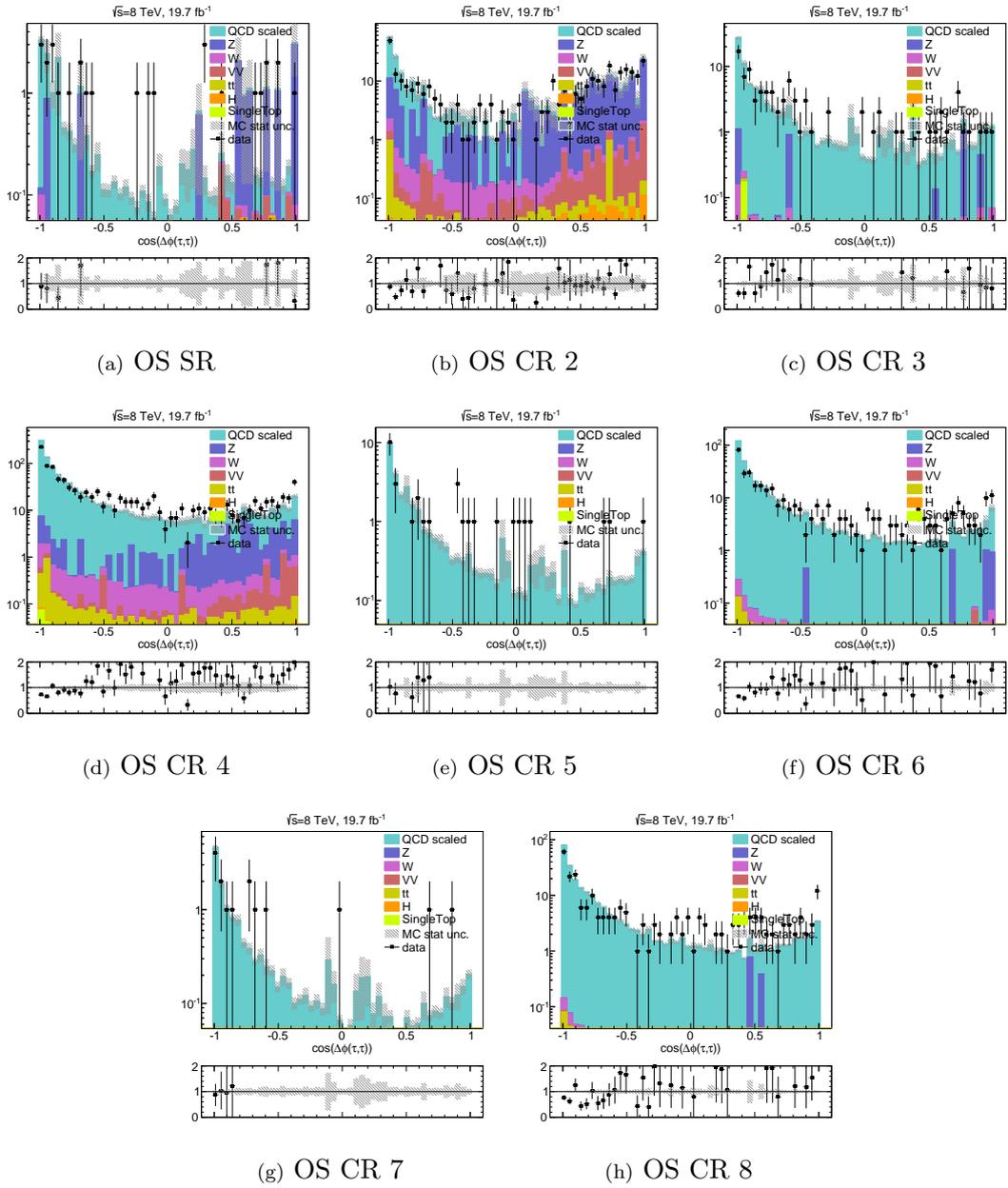
FIGURE B.23: Logarithmic cosine of the difference of the τ in ϕ for OS

B.4.6 $\Delta\eta(\tau,\tau)$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
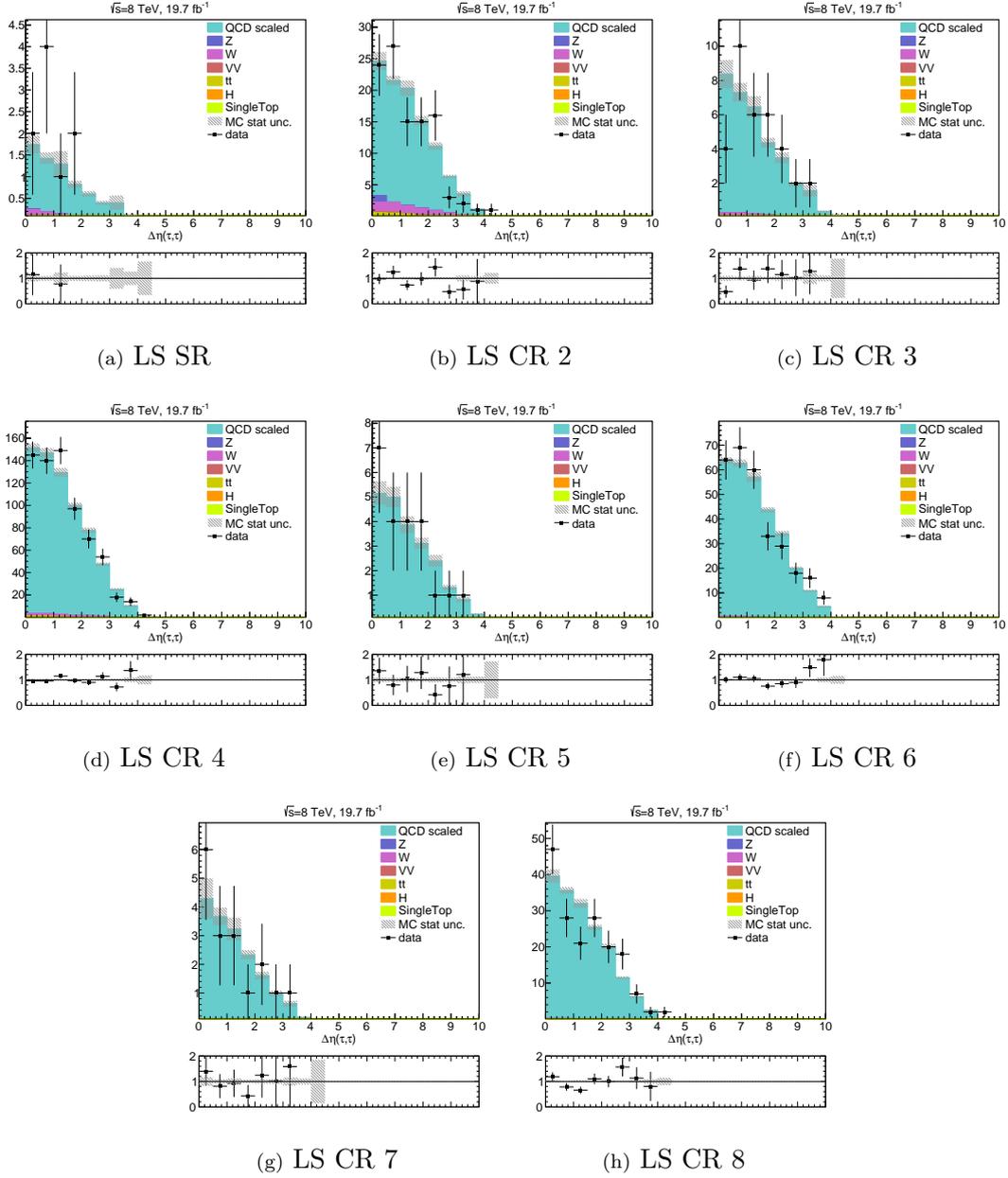
FIGURE B.24: Linear difference of the τ in η for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
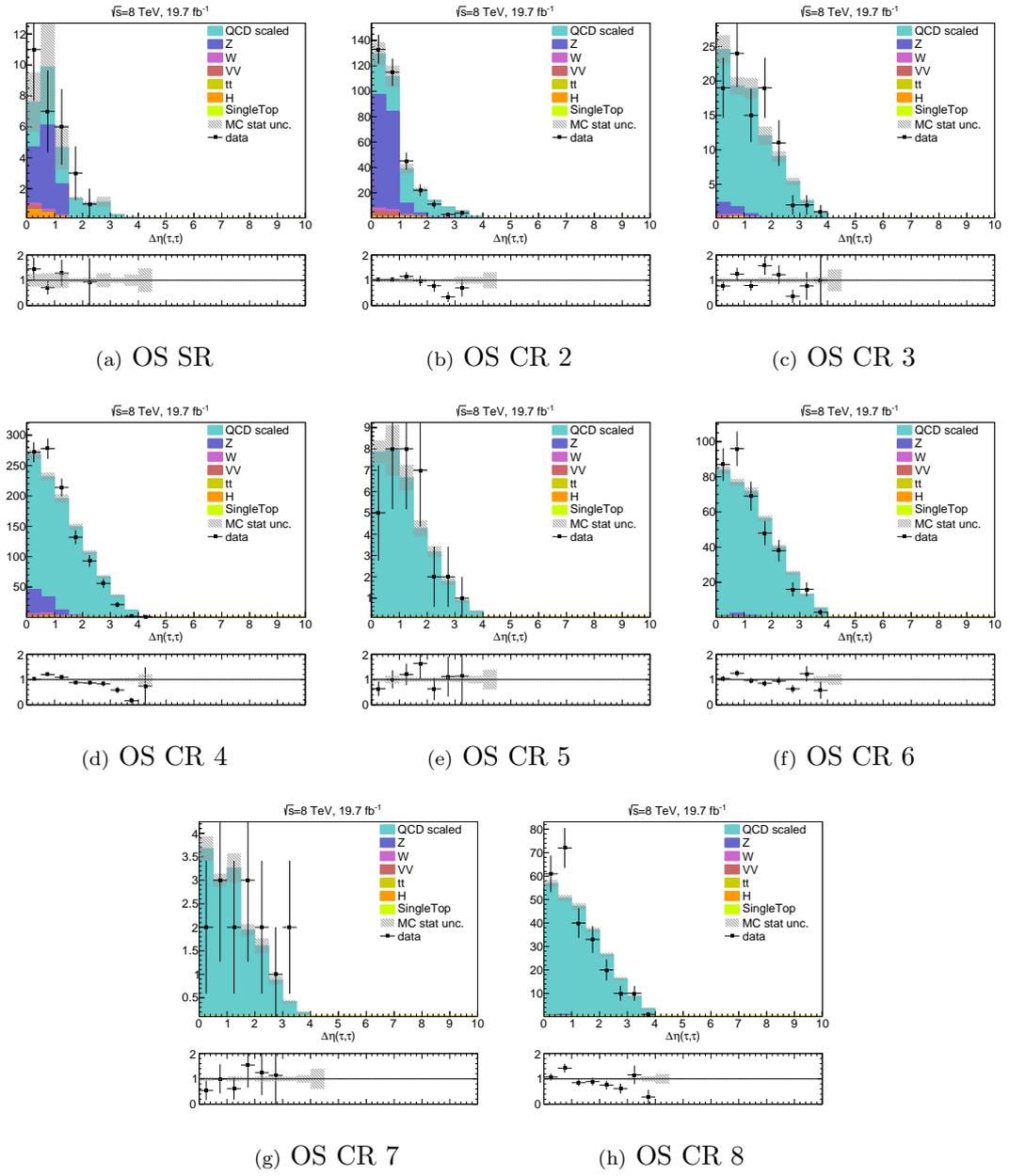
FIGURE B.25: Linear difference of the τ in η for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
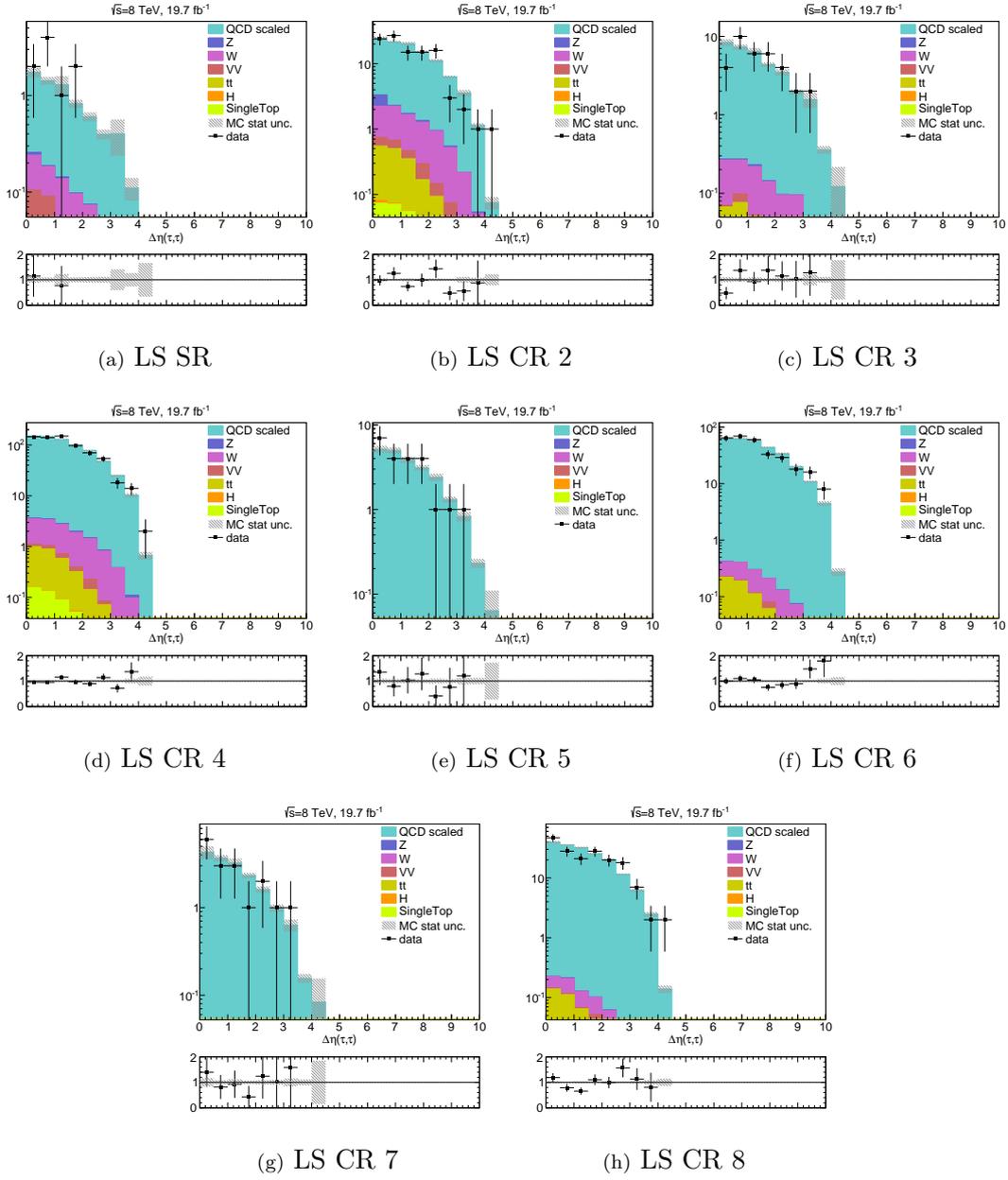
FIGURE B.26: Logarithmic difference of the τ in *eta* for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
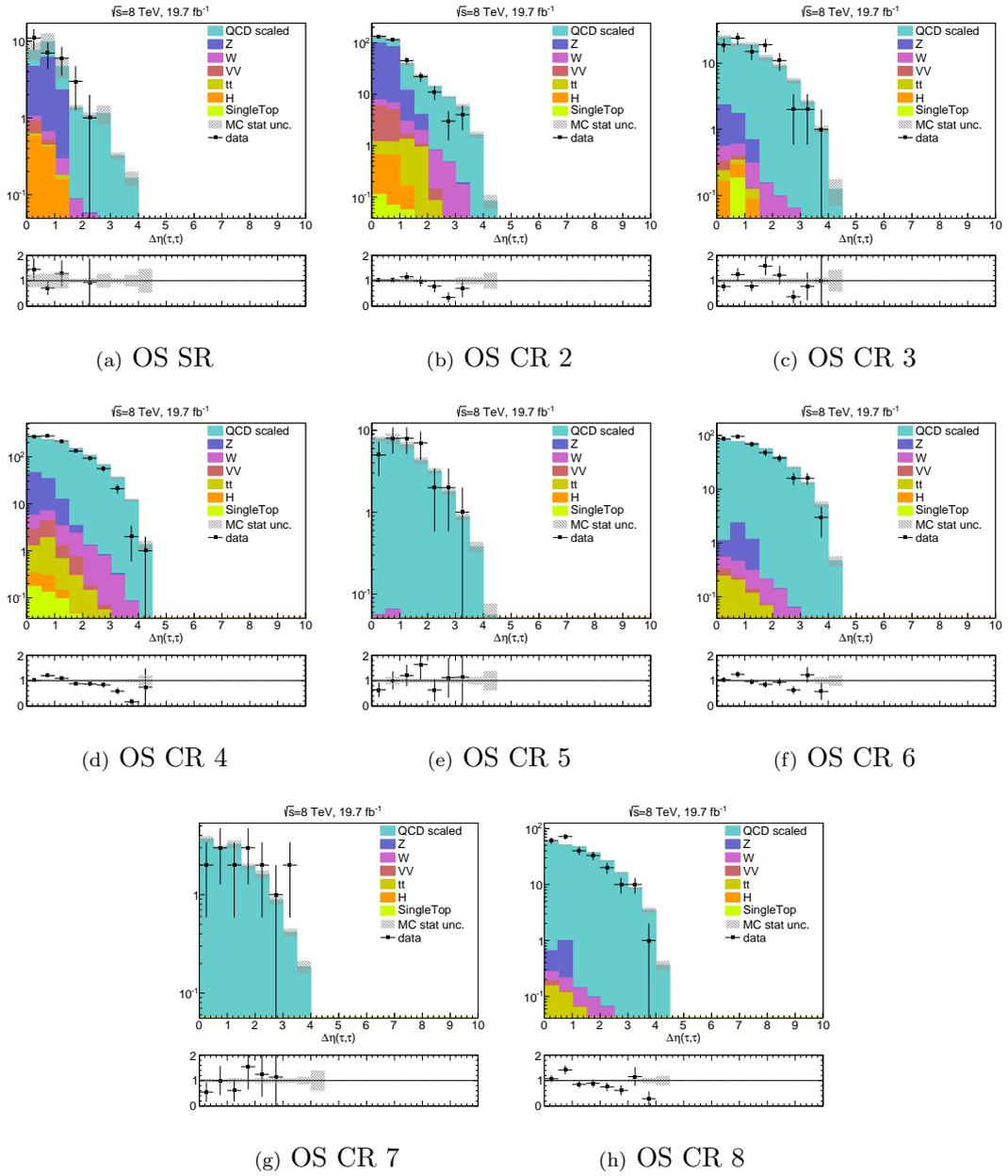
FIGURE B.27: Logarithmic difference of the τ in *eta* for OS

B.4.7 $M^{\tau,\tau}$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
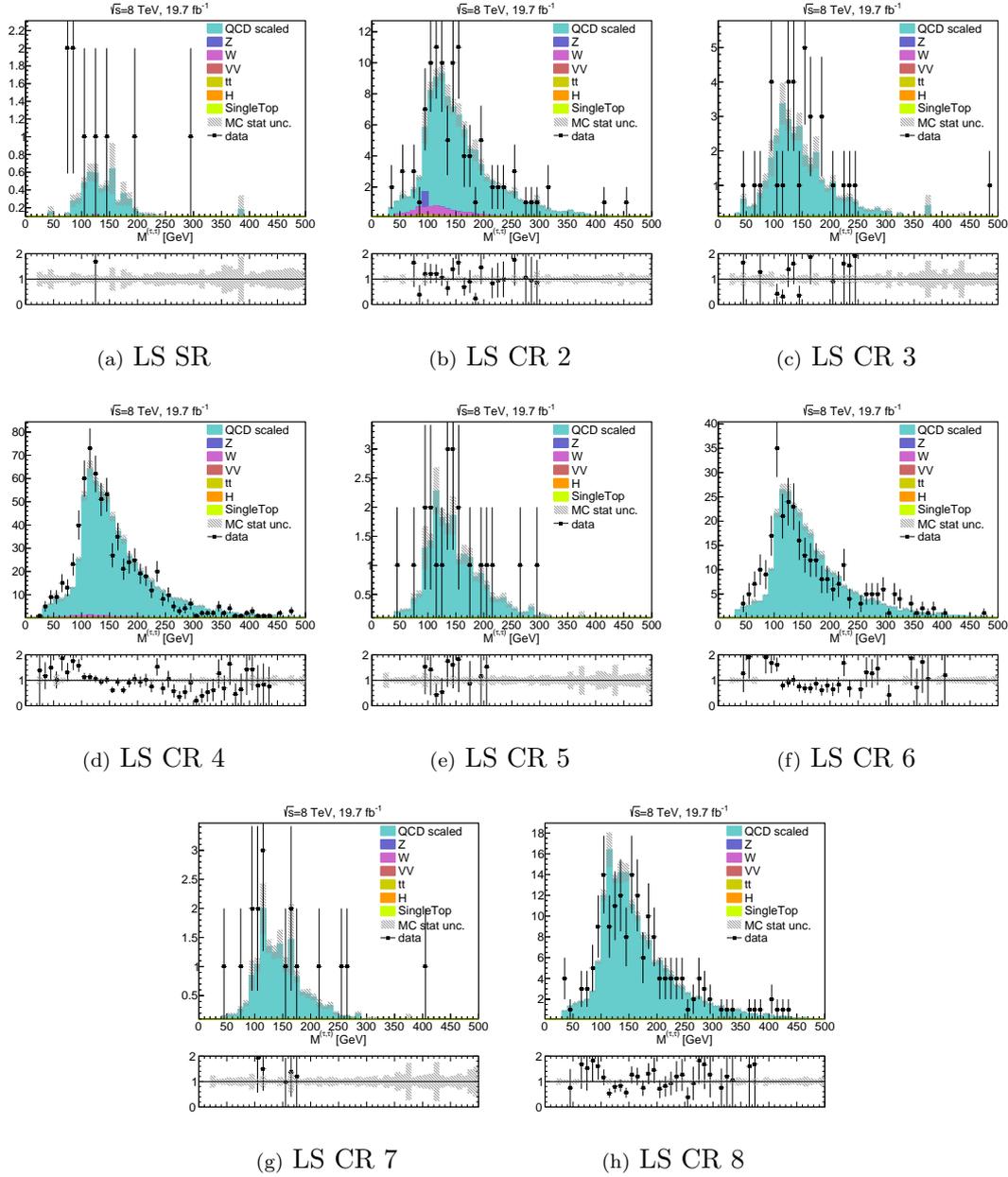
FIGURE B.28: Linear invariant mass of the di-τ-system for LS

(a) OS SR   (b) OS CR 2   (c) OS CR 3

(d) OS CR 4   (e) OS CR 5   (f) OS CR 6

(g) OS CR 7   (h) OS CR 8
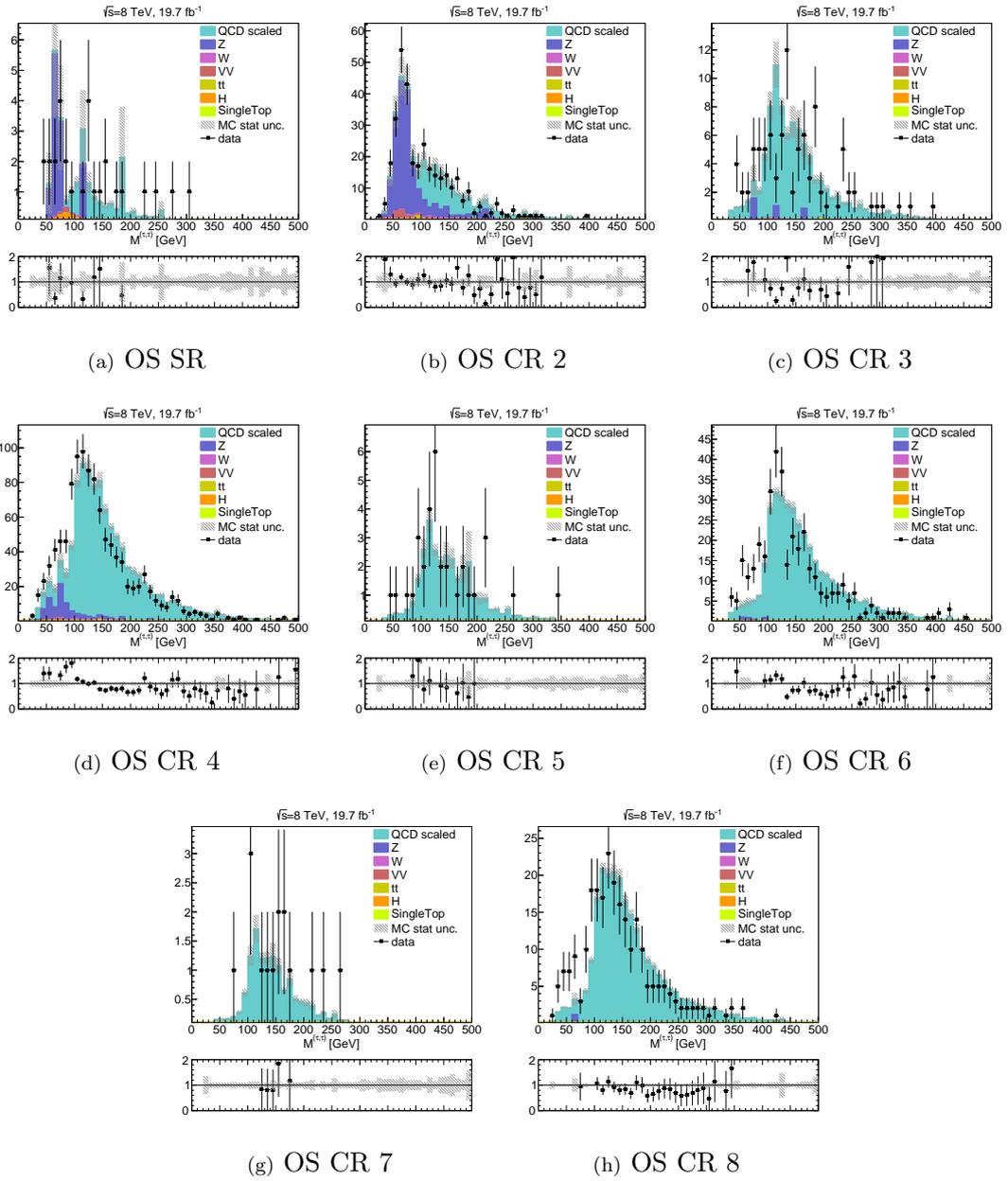
FIGURE B.29: Linear invariant mass of the di-τ-system for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
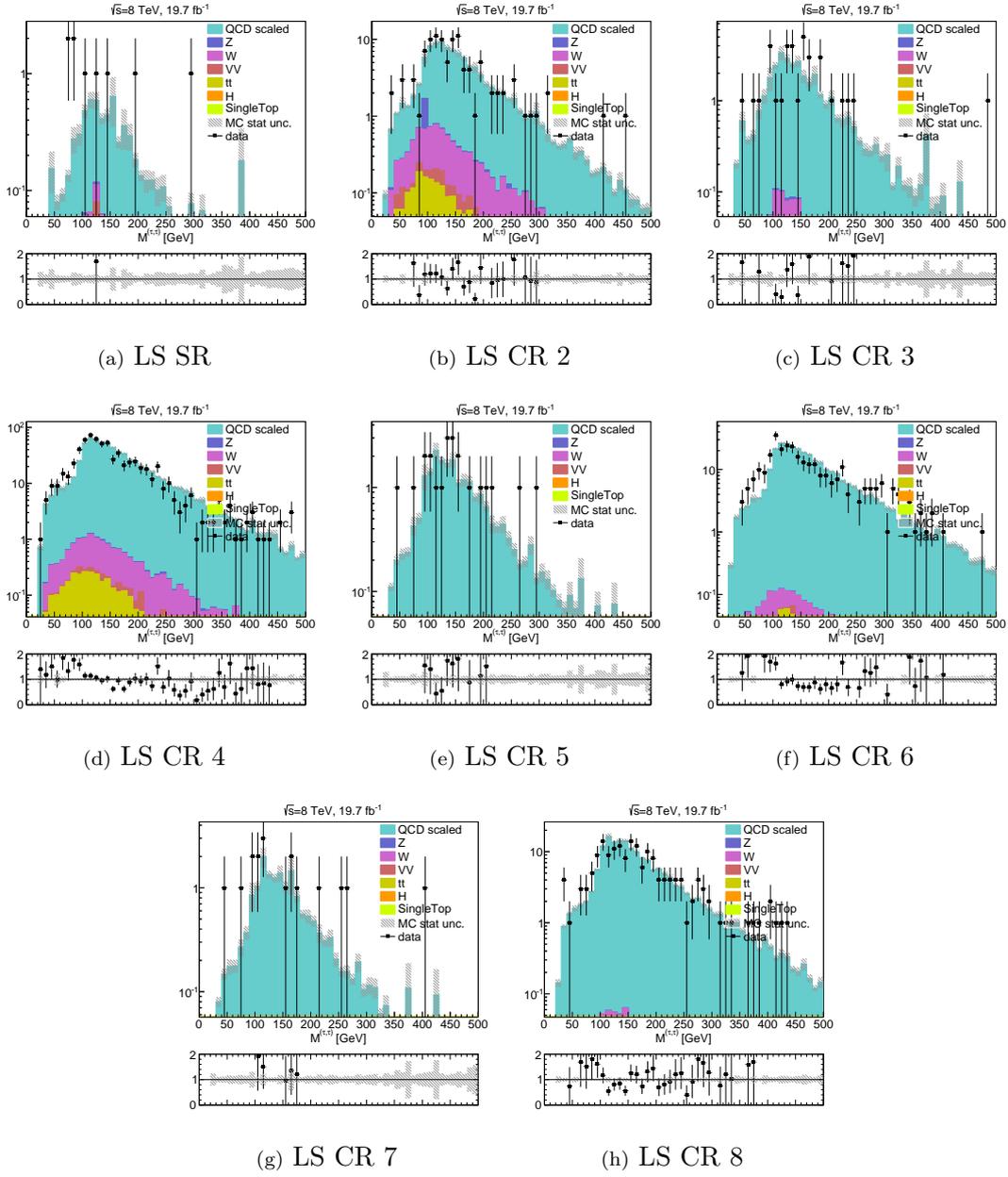
FIGURE B.30: Logarithmic invariant mass of the di-τ-system for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
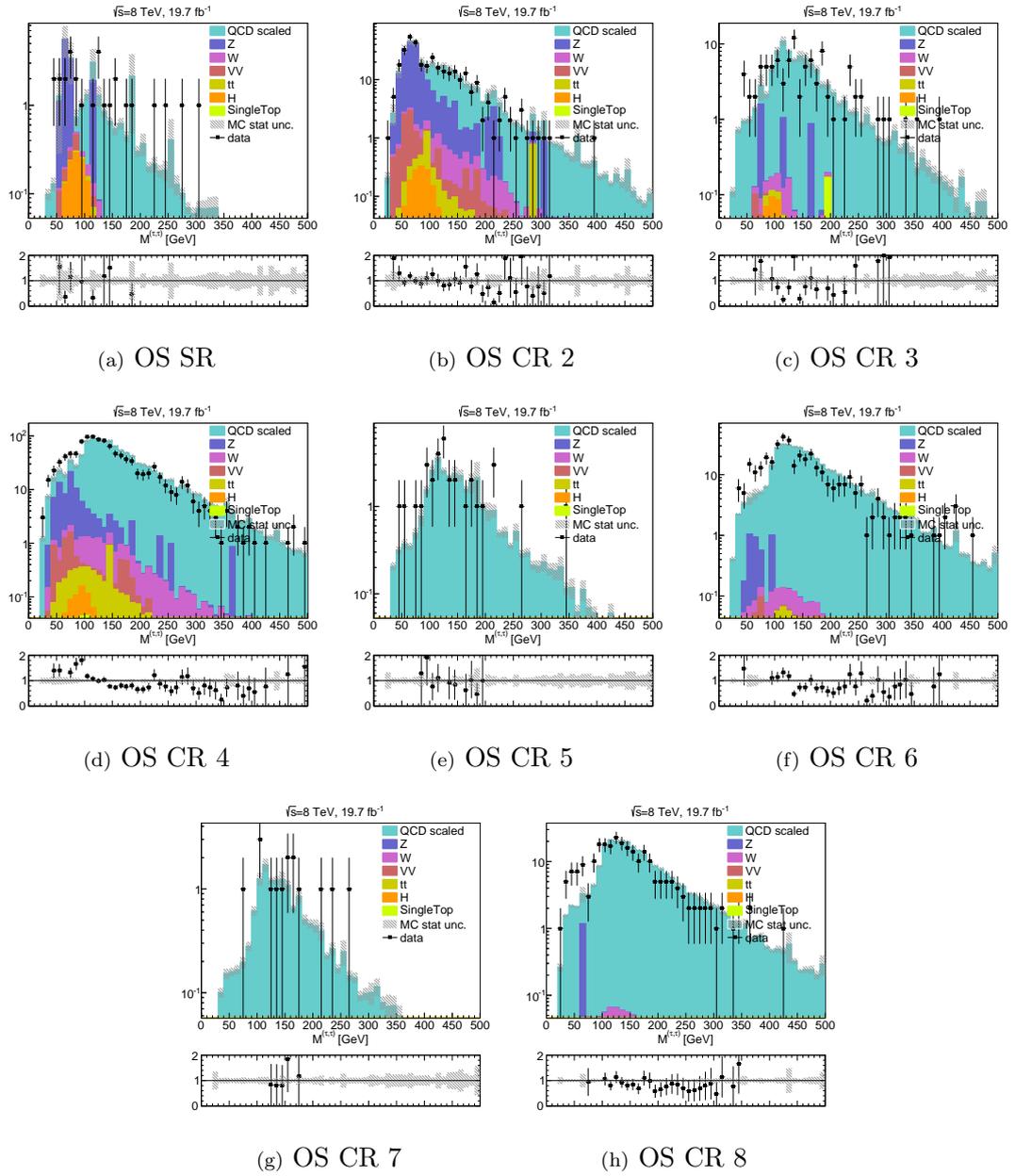
FIGURE B.31: Logarithmic invariant mass of the di-τ-system for OS

B.4.8 H_T

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
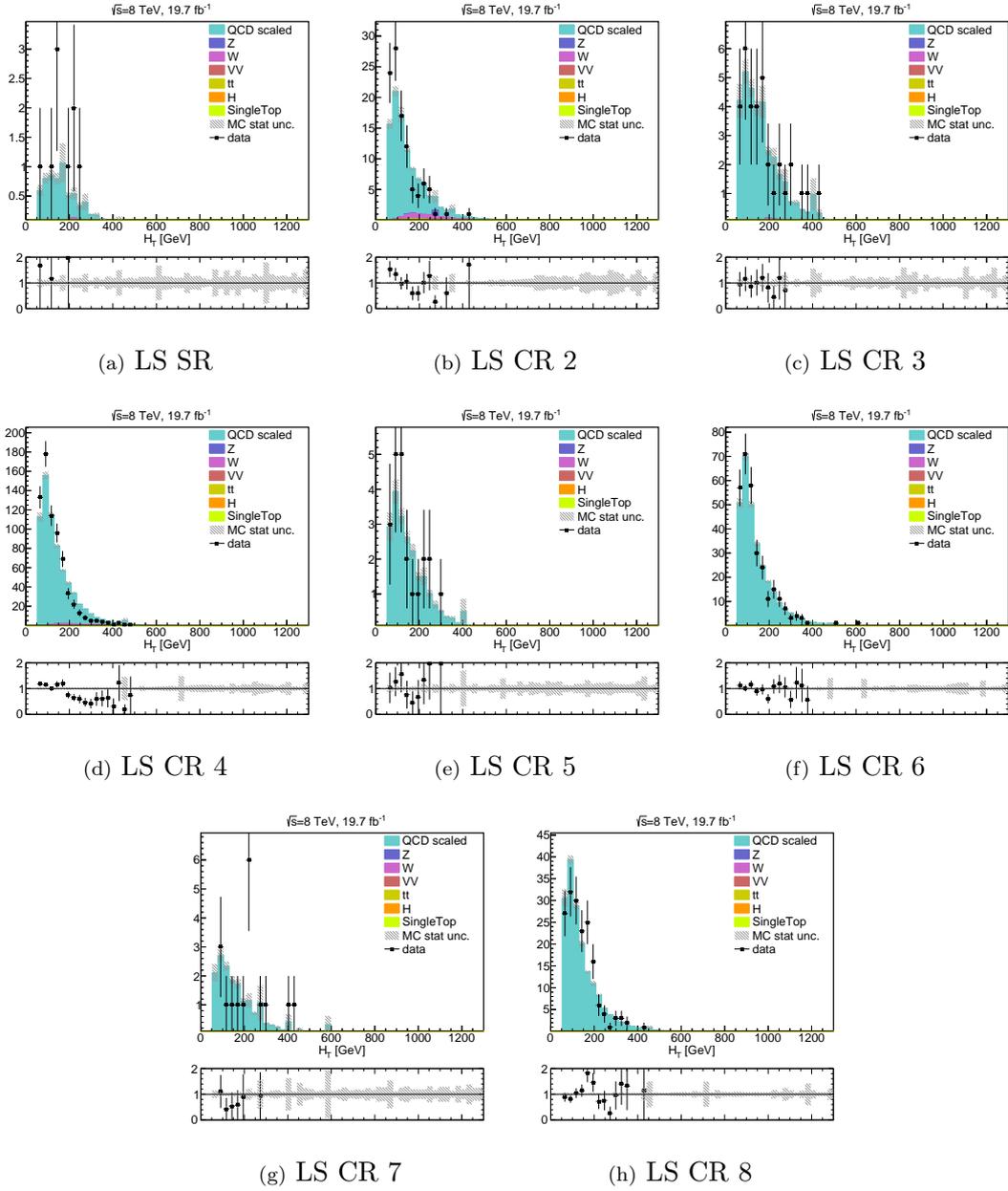
FIGURE B.32: Linear scalar $\sum p_T^{jet}$ for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
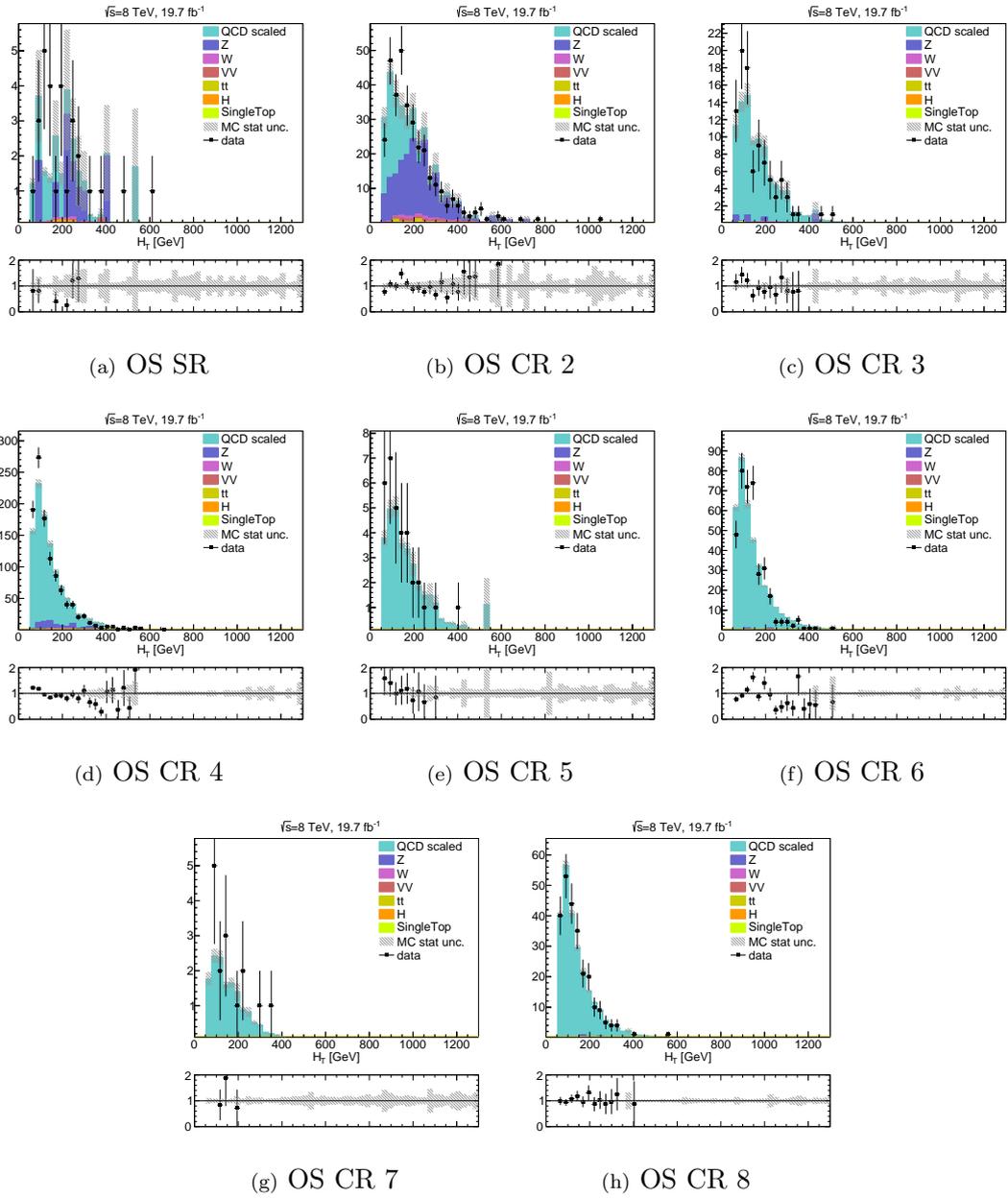
FIGURE B.33: Linear scalar $\sum p_T^{jet}$ for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
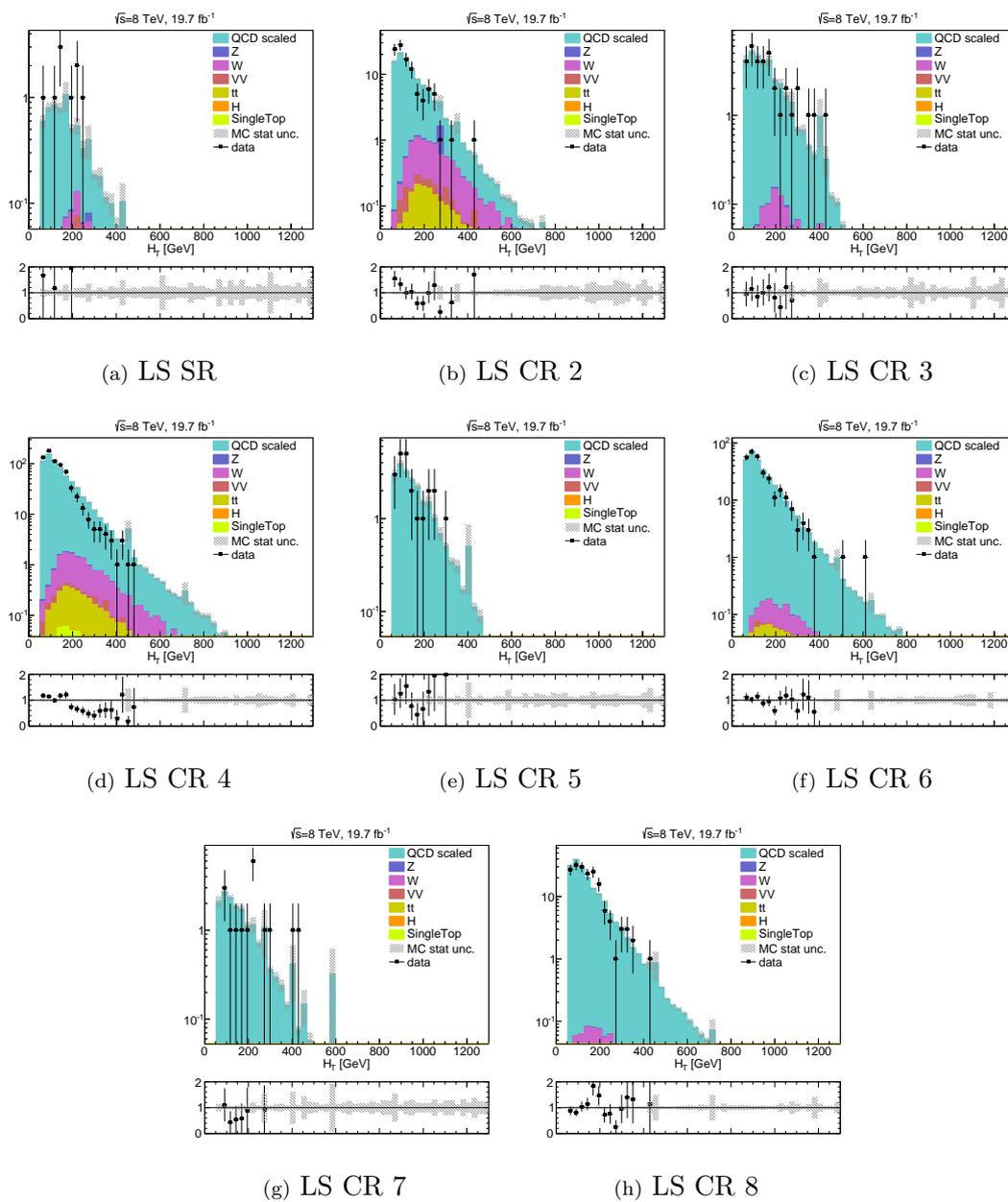
FIGURE B.34: Logarithmic scalar $\sum p_T^{jet}$ for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
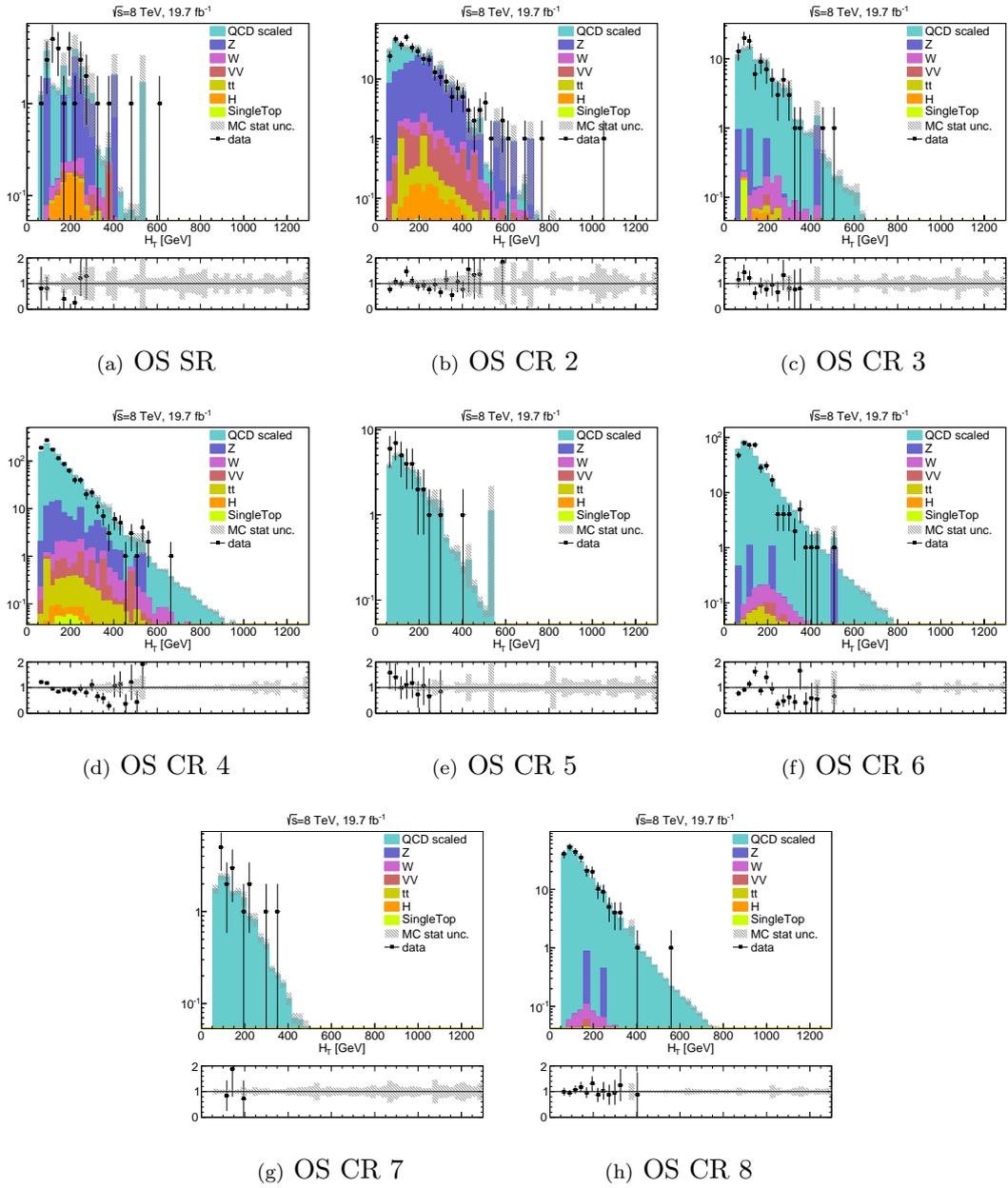
FIGURE B.35: Logarithmic scalar $\sum p_T^{jet}$ for OS

B.4.9 $H_T + \sum p_T^\tau$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
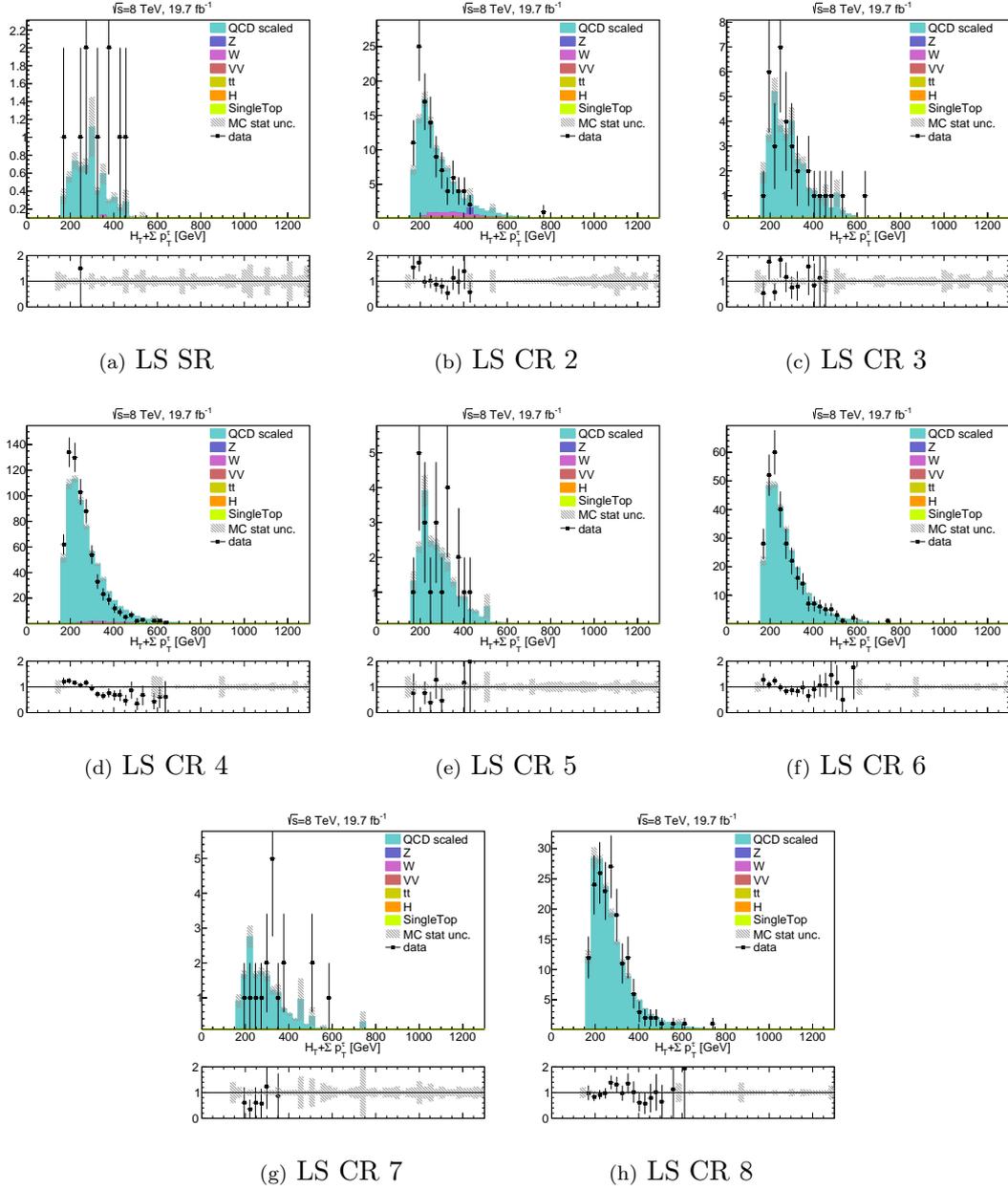
FIGURE B.36: Linear scalar $\sum p_T^{jet} + \sum p_T^\tau$ for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
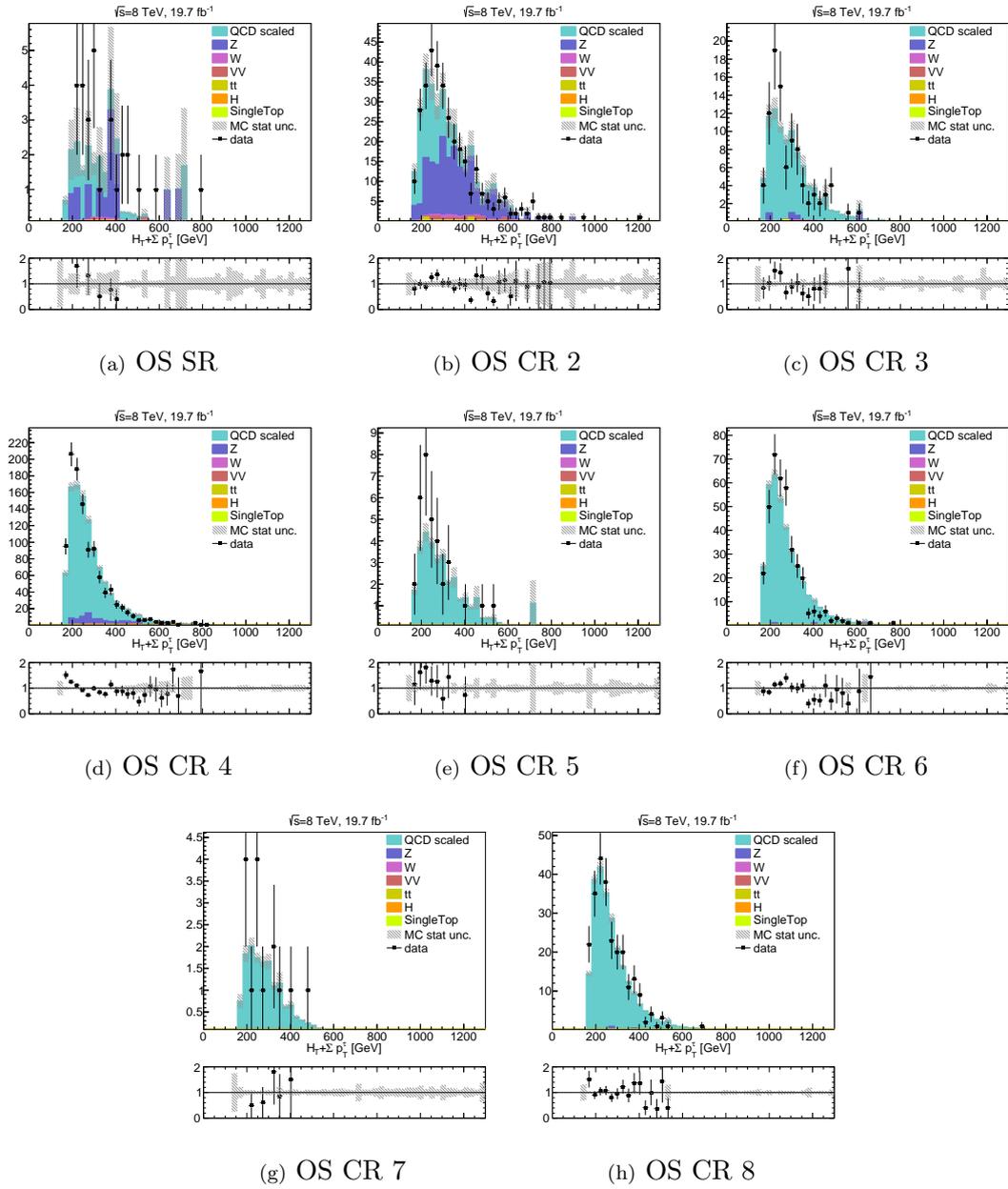
FIGURE B.37: Linear scalar $\sum p_T^{jet} + \sum p_T^{\tau}$ for OS

(a) LS SR (b) LS CR 2 (c) LS CR 3

(d) LS CR 4 (e) LS CR 5 (f) LS CR 6

(g) LS CR 7 (h) LS CR 8
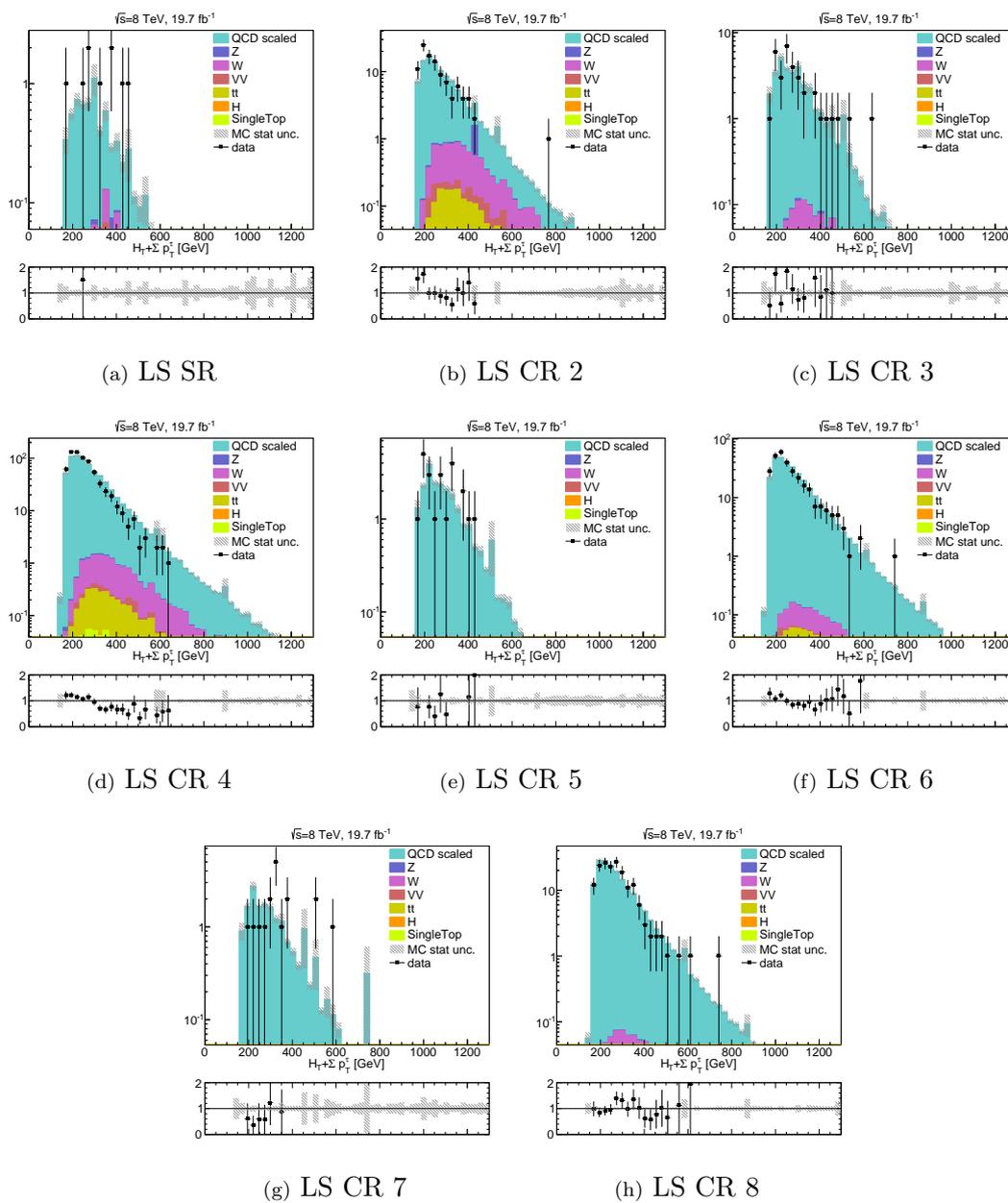
FIGURE B.38: Logarithmic scalar $\sum p_T^{jet} + \sum p_T^{\tau}$ for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
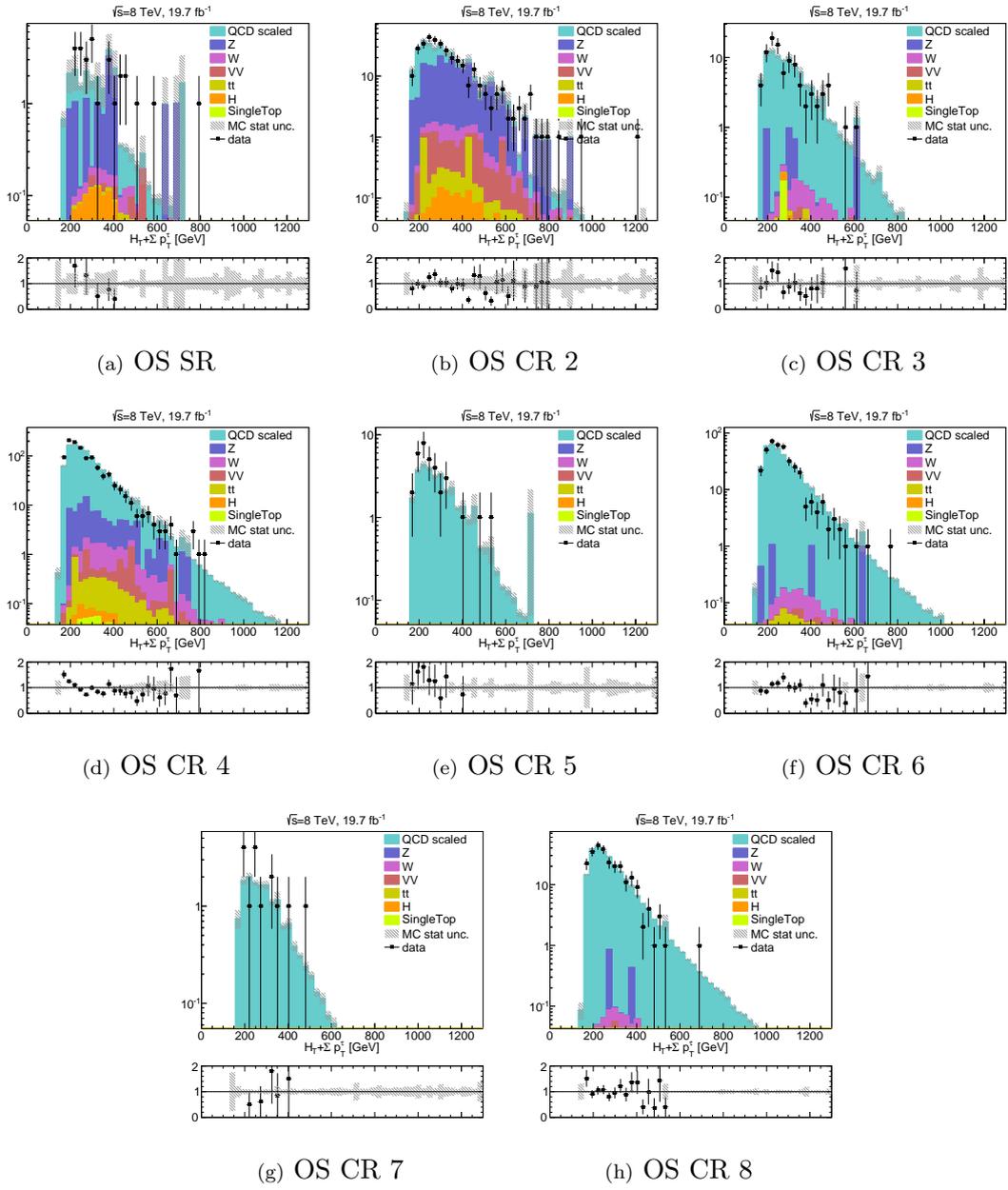
FIGURE B.39: Logarithmic scalar $\sum p_T^{jet} + \sum p_T^{\tau}$ for OS

B.4.10 $p_T^{jet\ 1}$

(a) LS SR (b) LS CR 2 (c) LS CR 3

(d) LS CR 4 (e) LS CR 5 (f) LS CR 6

(g) LS CR 7 (h) LS CR 8
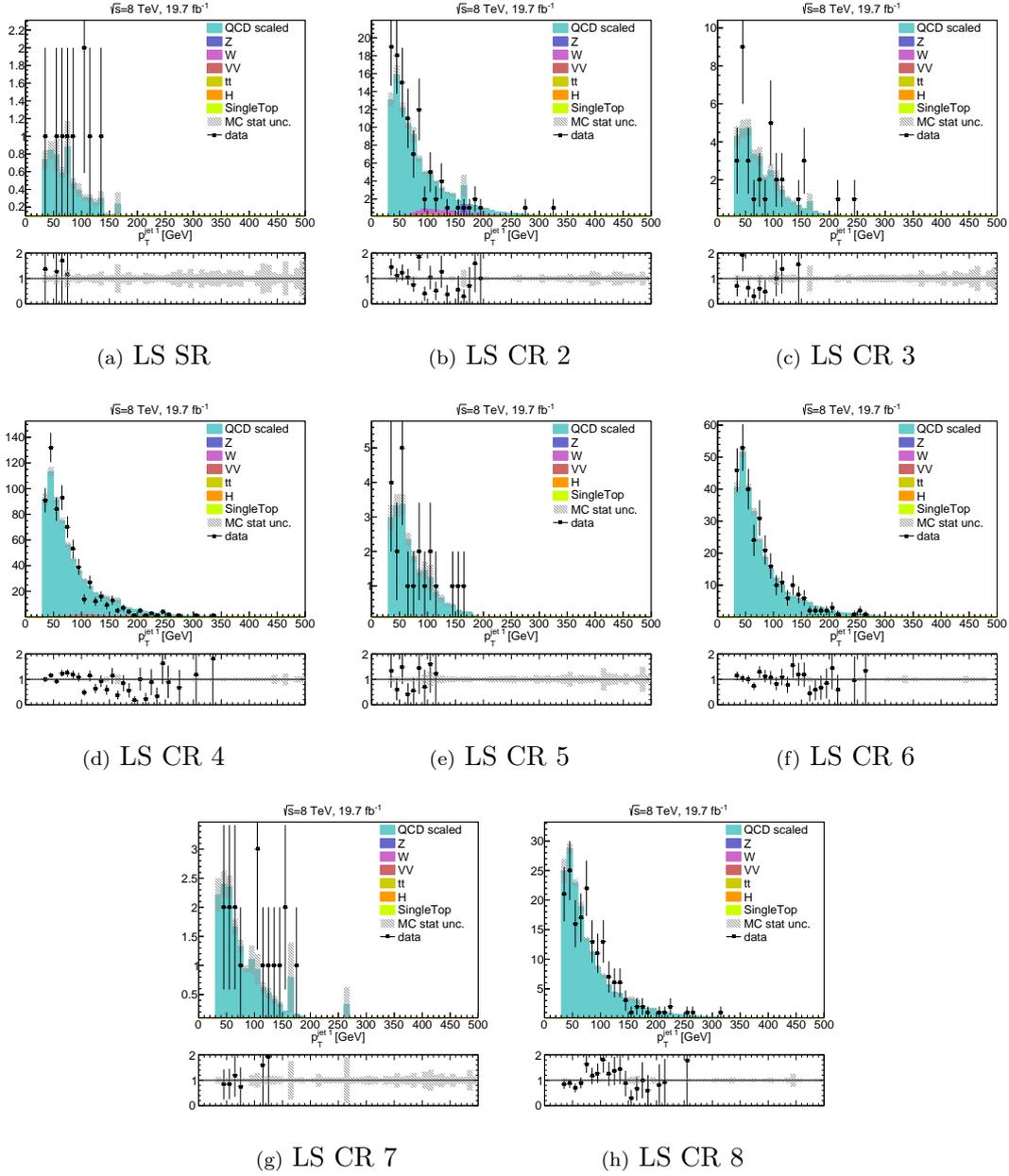
FIGURE B.40: Linear transverse momentum of the leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
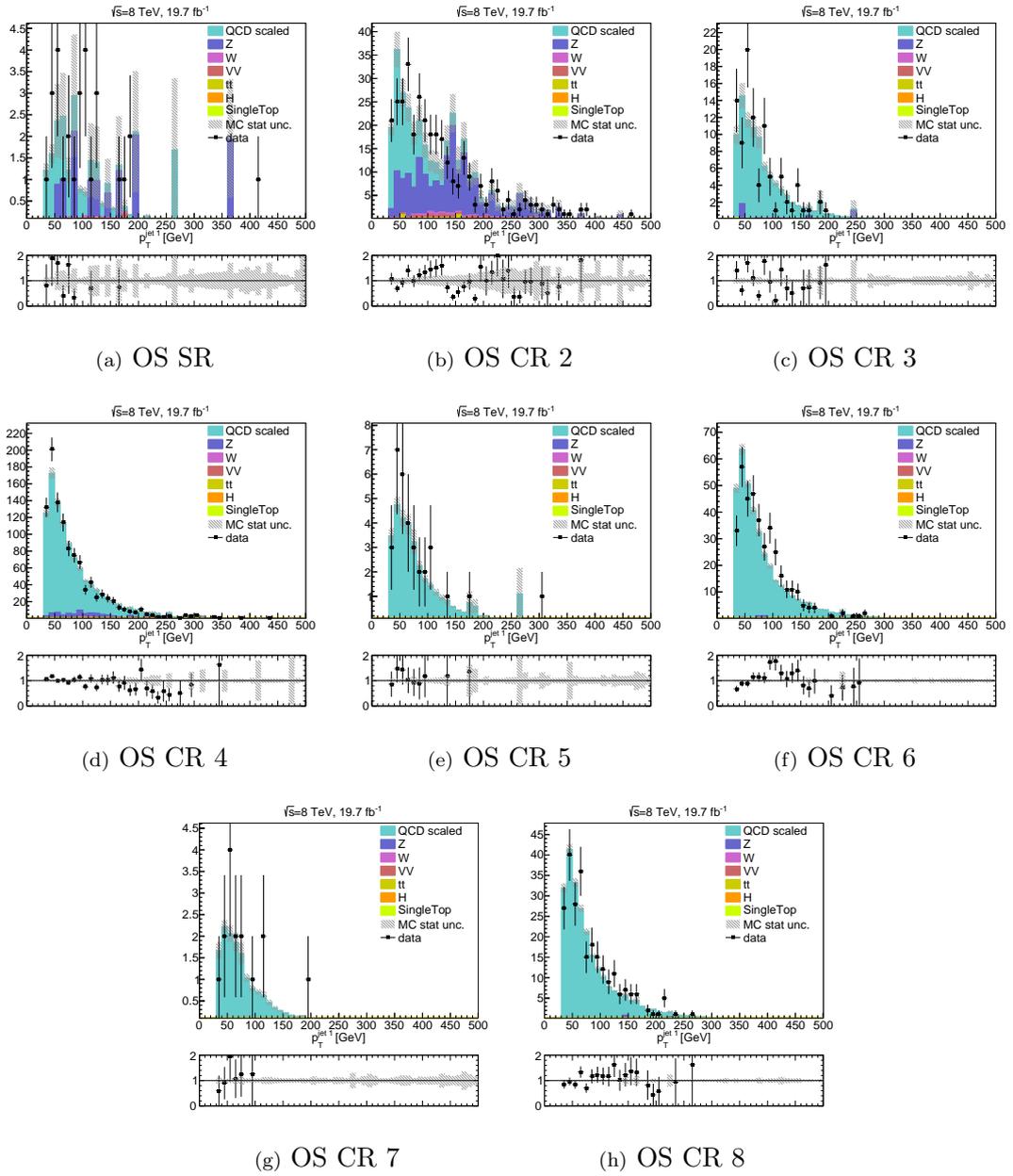
FIGURE B.41: Linear transverse momentum of the leading jet for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
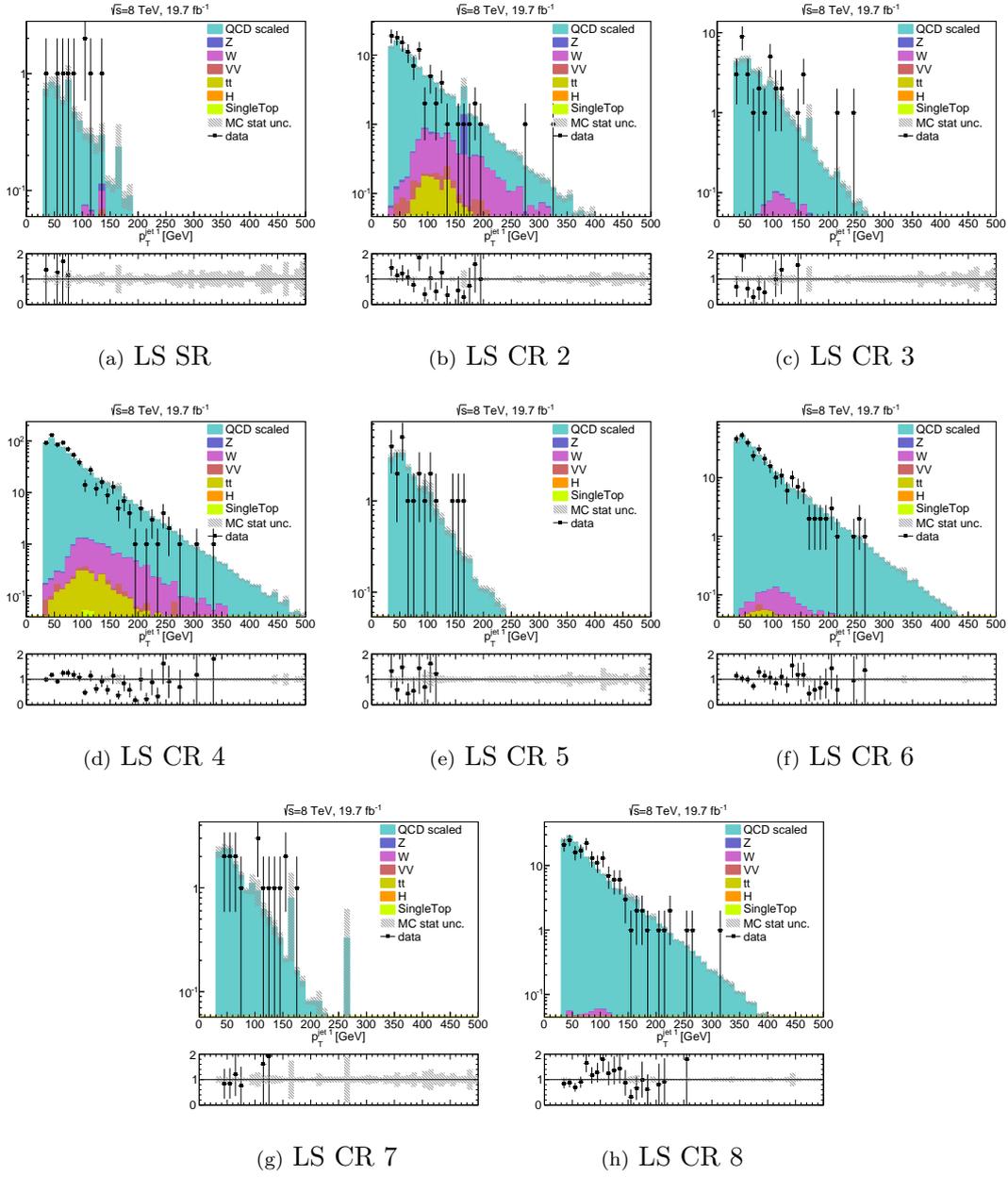
FIGURE B.42: Logarithmic transverse momentum of the leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
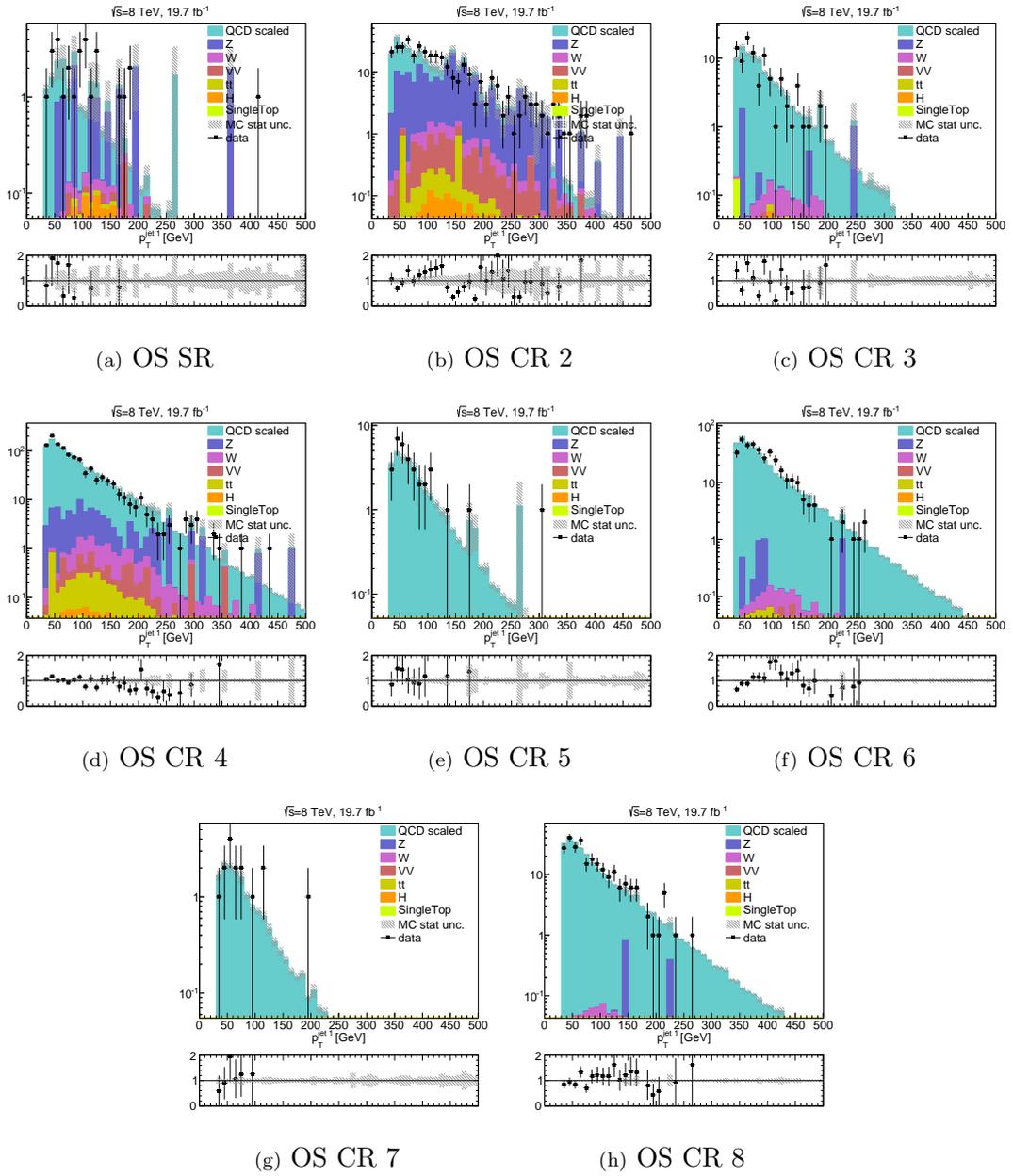
FIGURE B.43: Logarithmic transverse momentum of the leading jet for OS

B.4.11 $p_T^{jet\ 2}$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
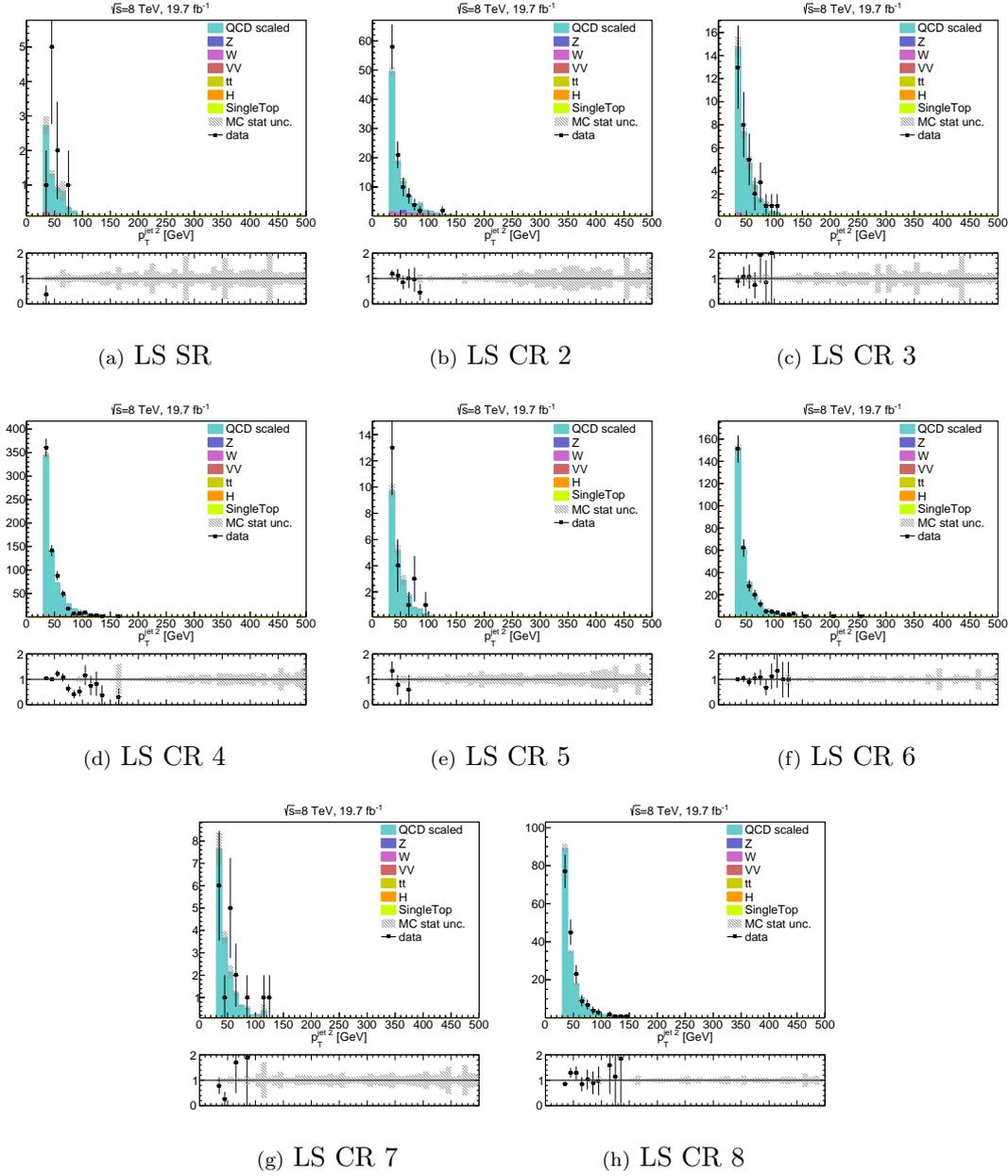
FIGURE B.44: Linear transverse momentum of the second to leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
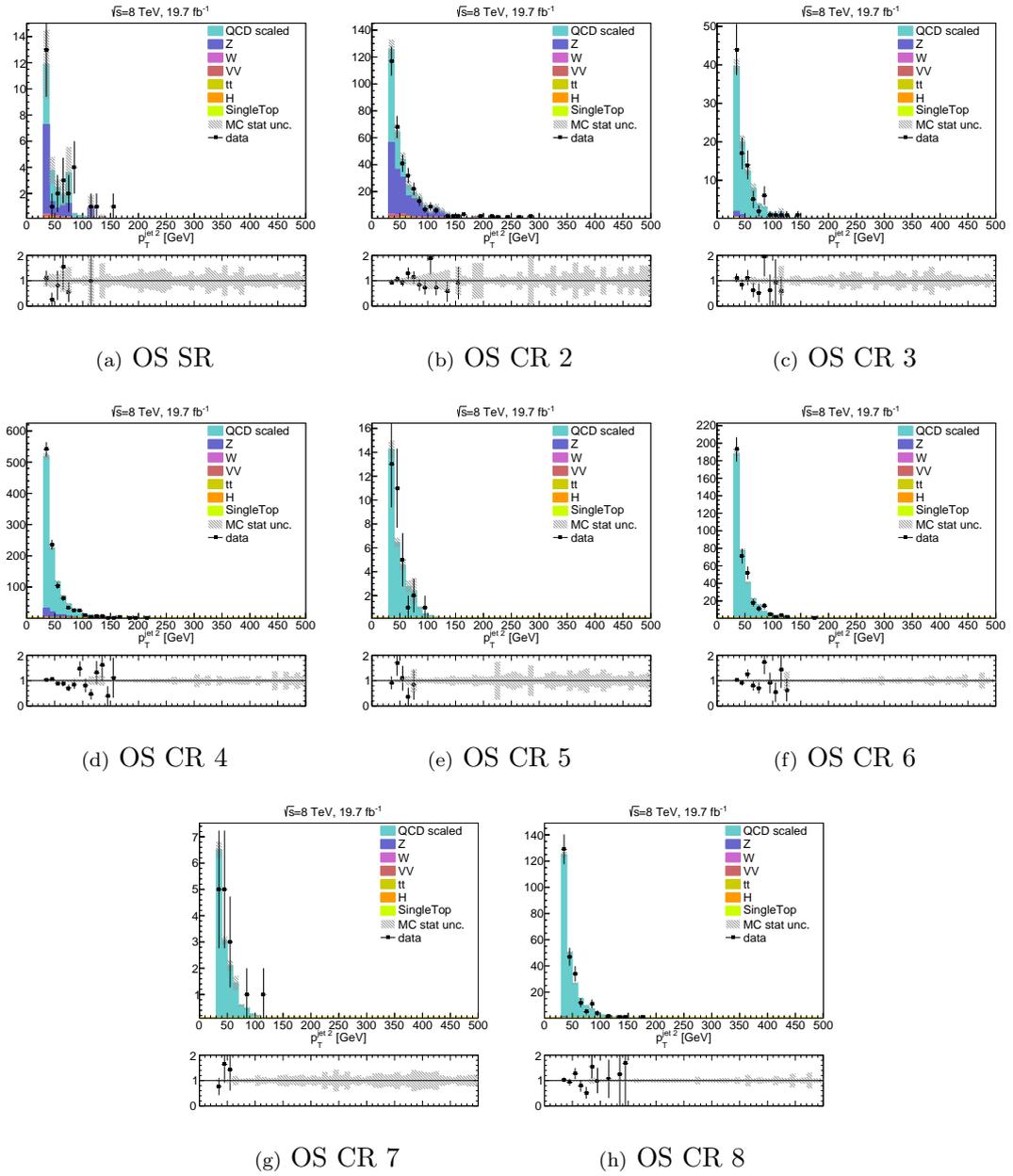
FIGURE B.45: Linear transverse momentum of the second to leading jet for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
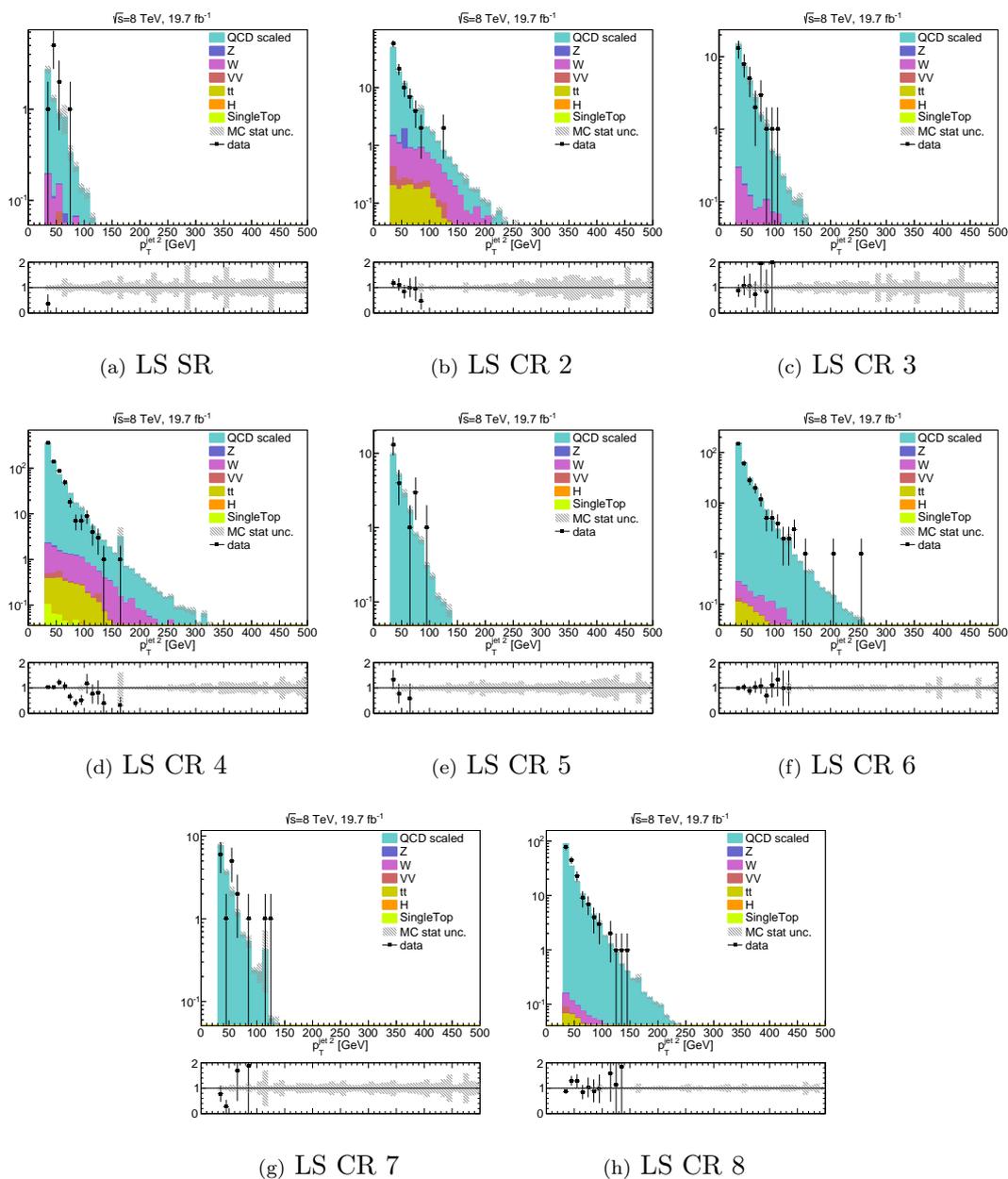
FIGURE B.46: Logarithmic transverse momentum of the second to leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
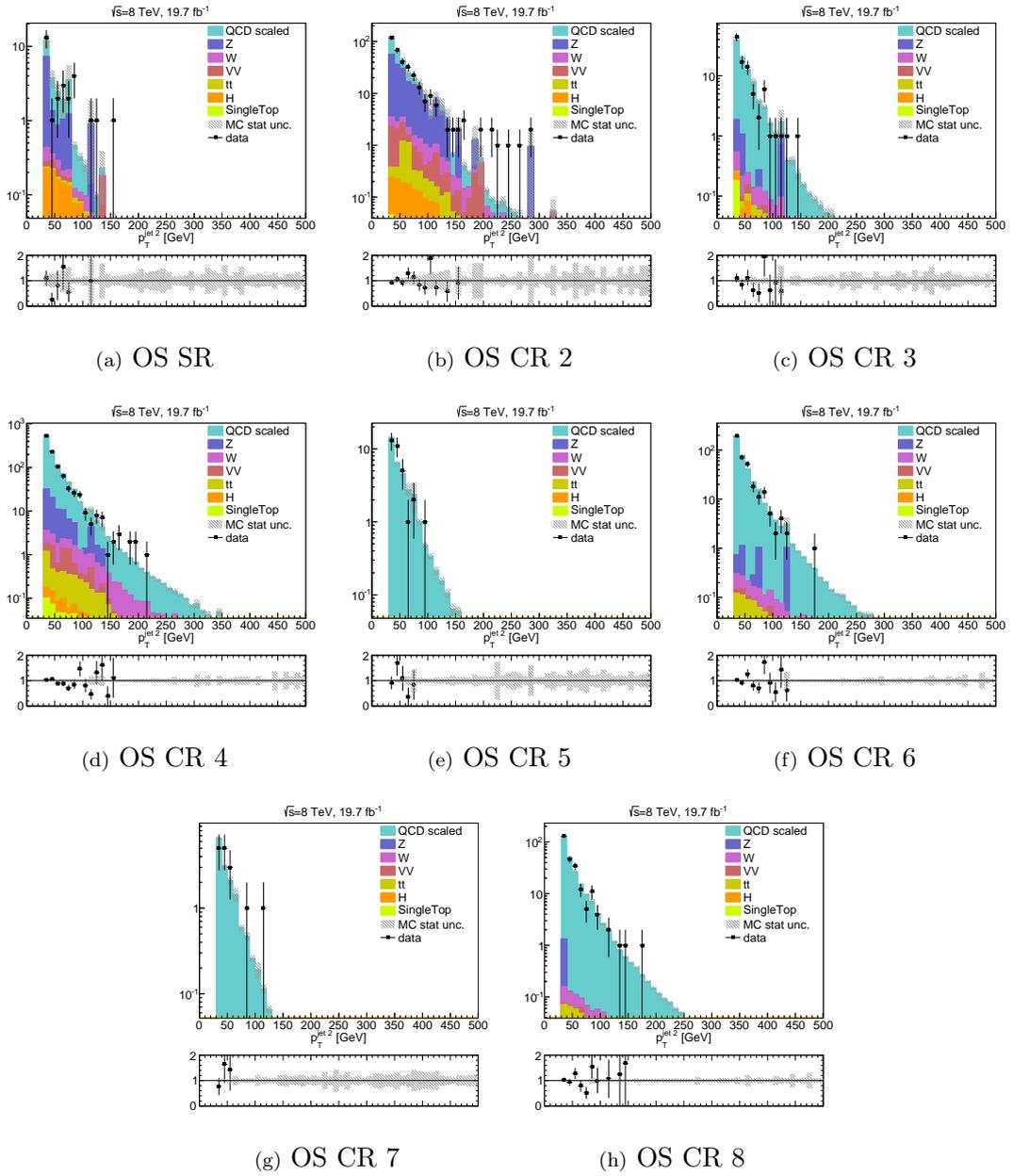
FIGURE B.47: Logarithmic transverse momentum of the second to leading jet for OS

B.4.12 Inclusive p_T^{jet}

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
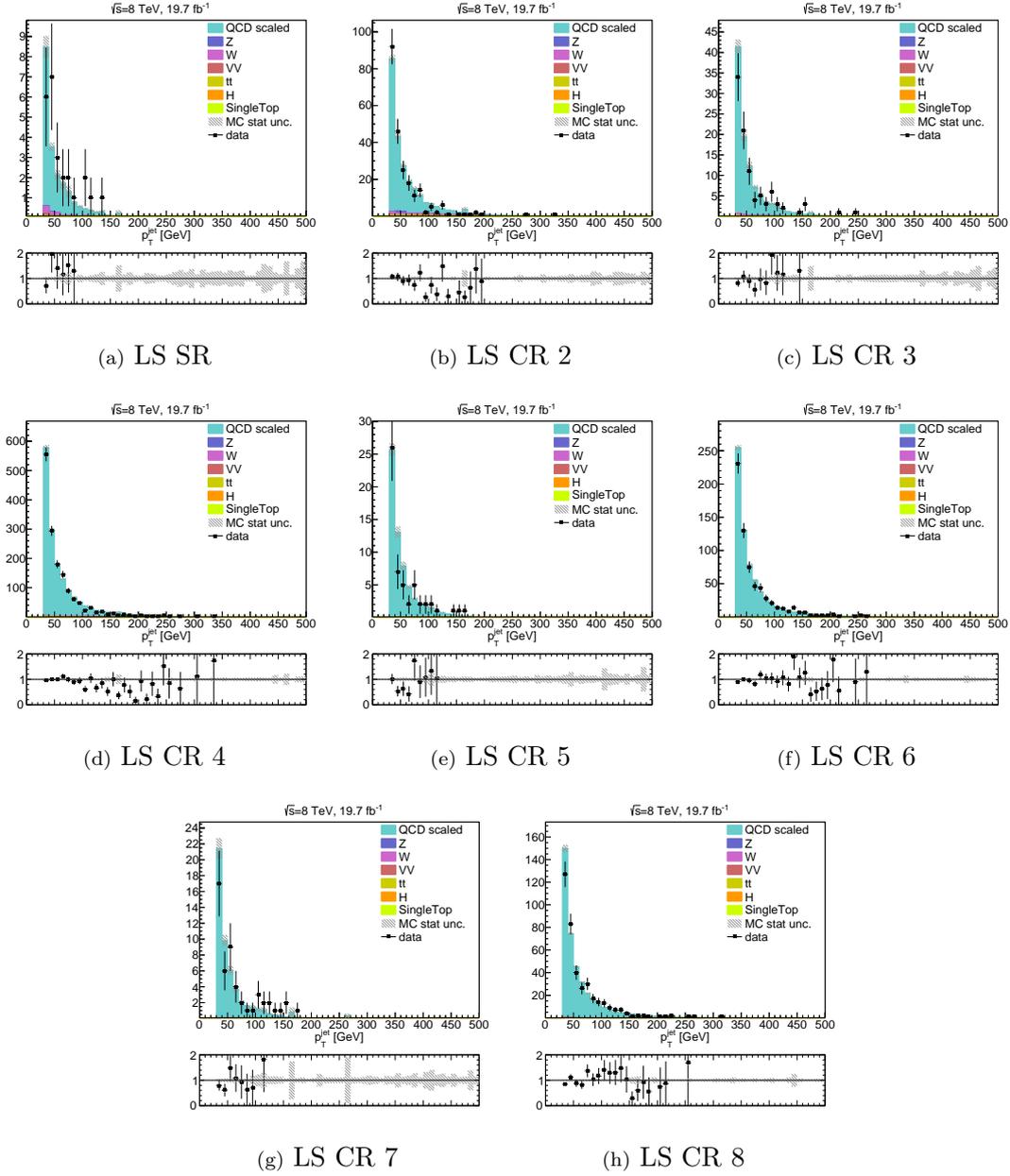
FIGURE B.48: Linear transverse momentum of all jets for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
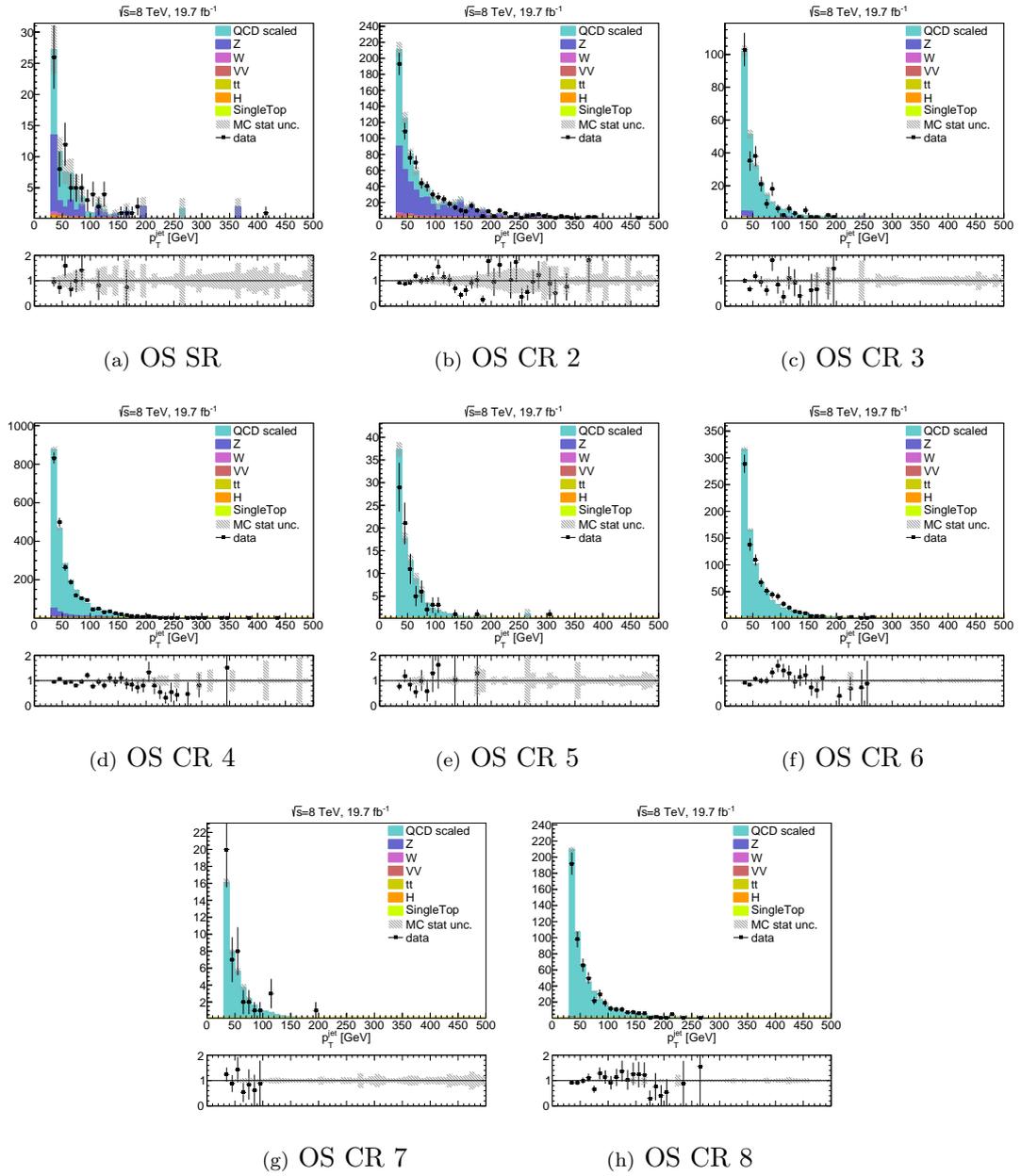
FIGURE B.49: Linear transverse momentum of all jets for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
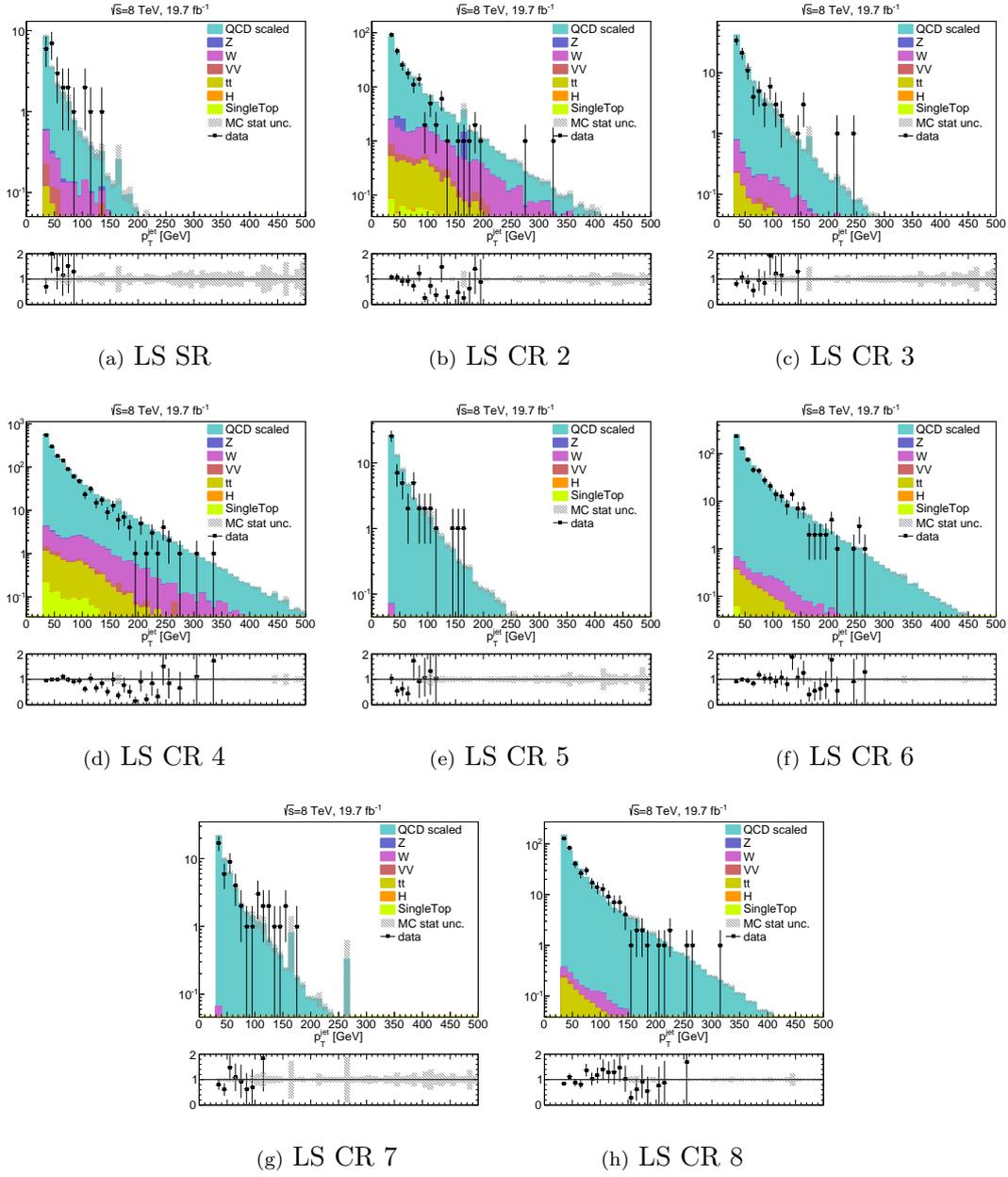
FIGURE B.50: Logarithmic transverse momentum of all jets for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
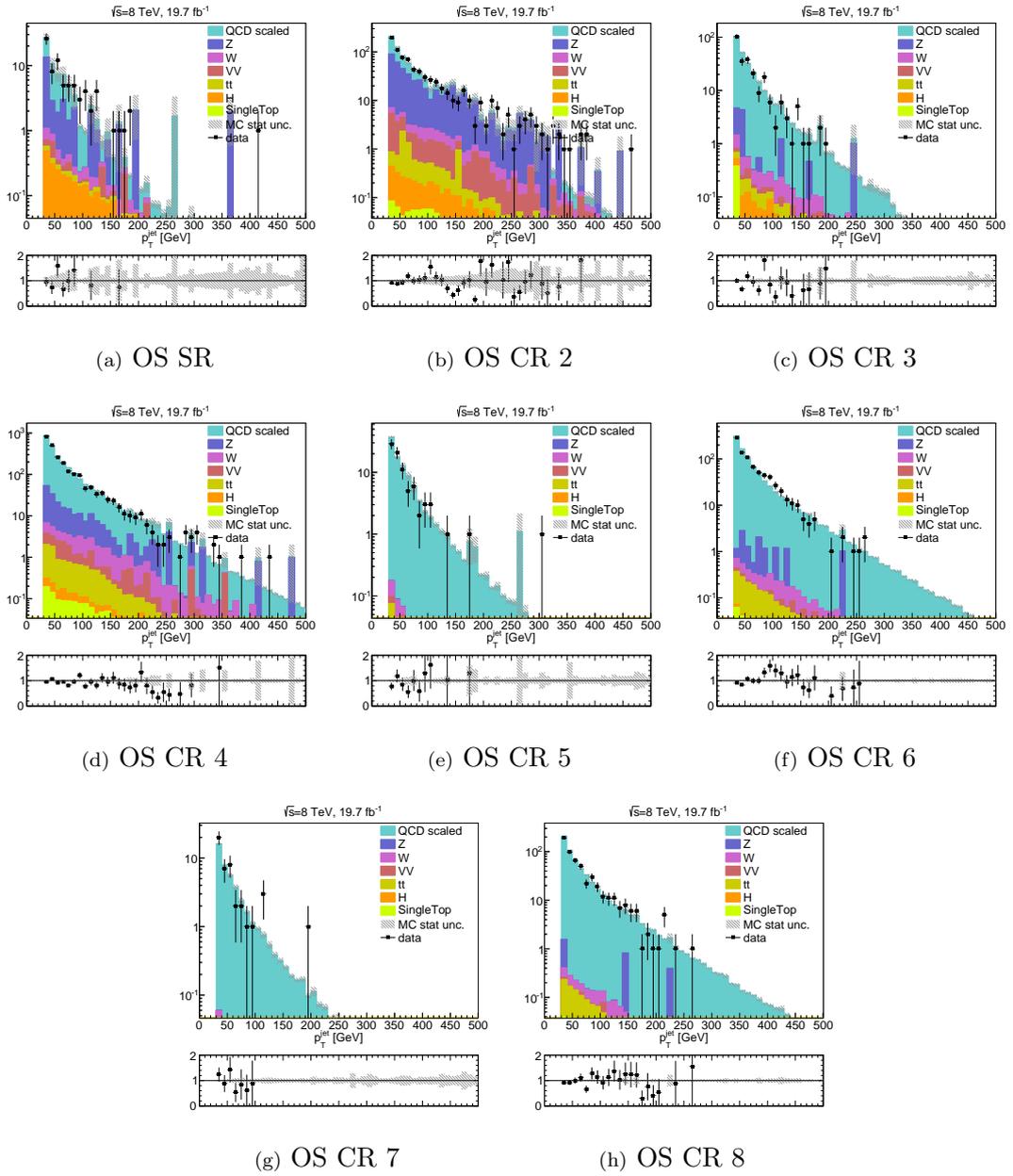
FIGURE B.51: Logarithmic transverse momentum of all jets for OS

B.4.13 $\eta^{jet\ 1}$

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
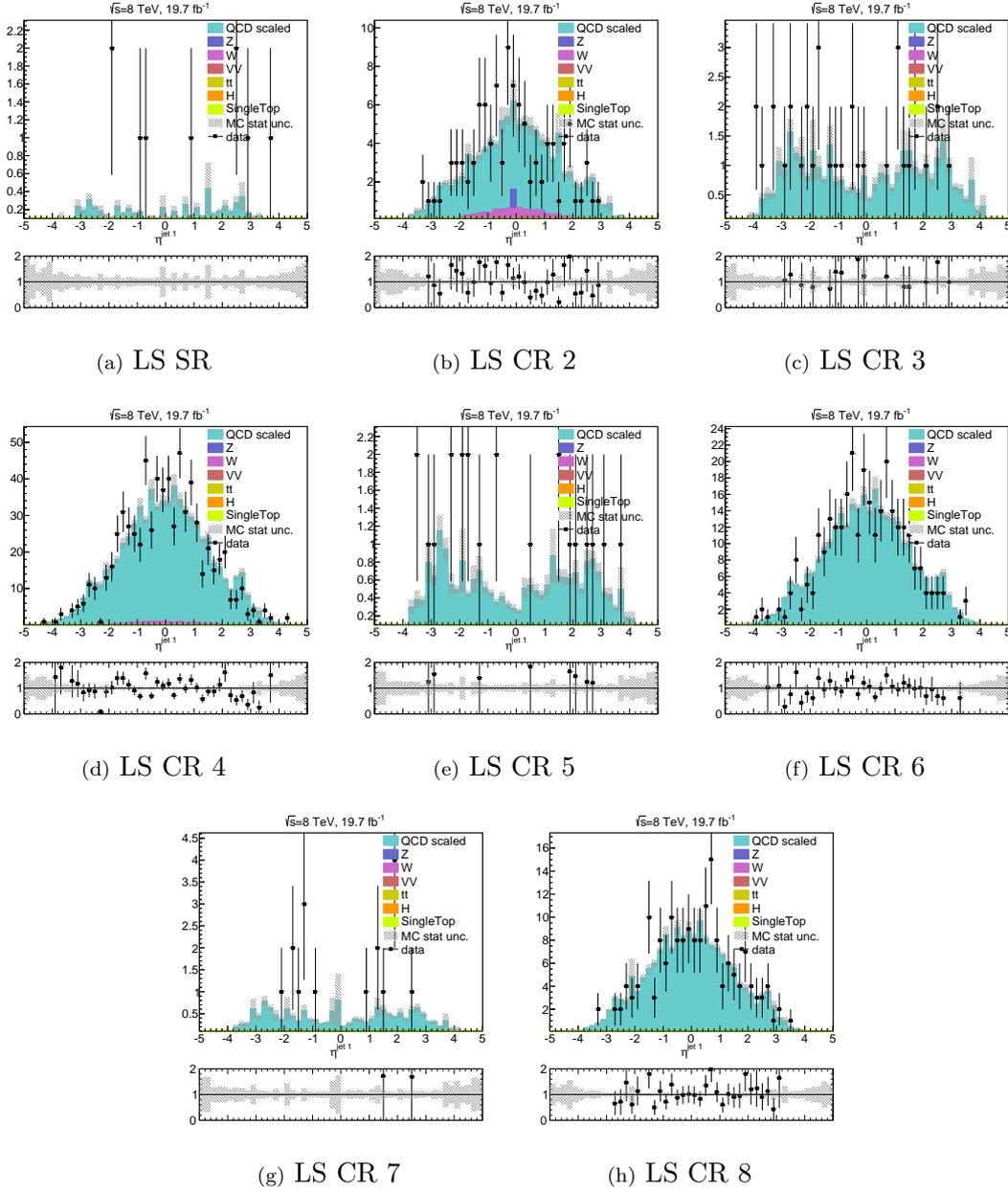
FIGURE B.52: Linear pseudorapidity of the leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
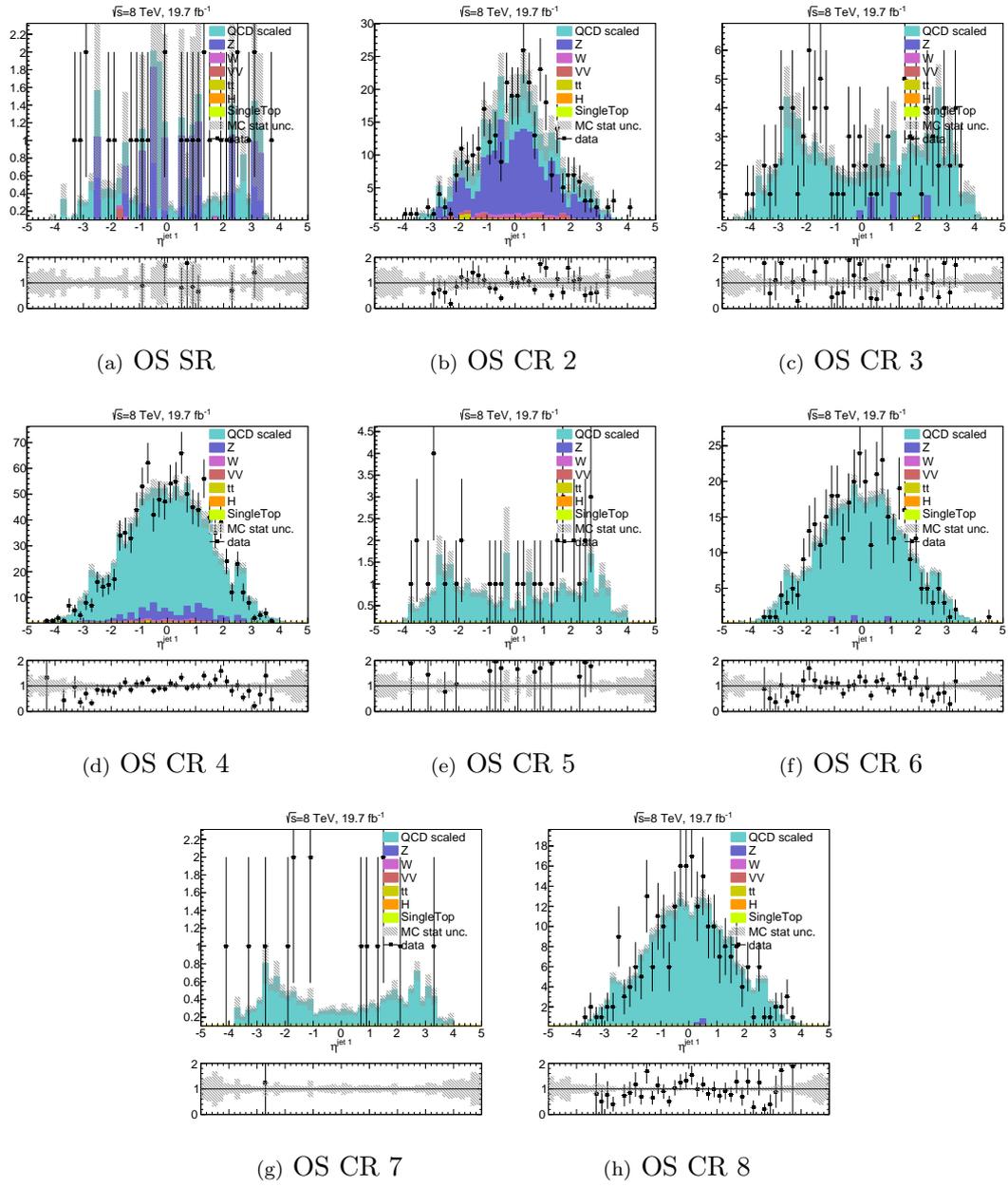
FIGURE B.53: Linear pseudorapidity of the leading jet for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
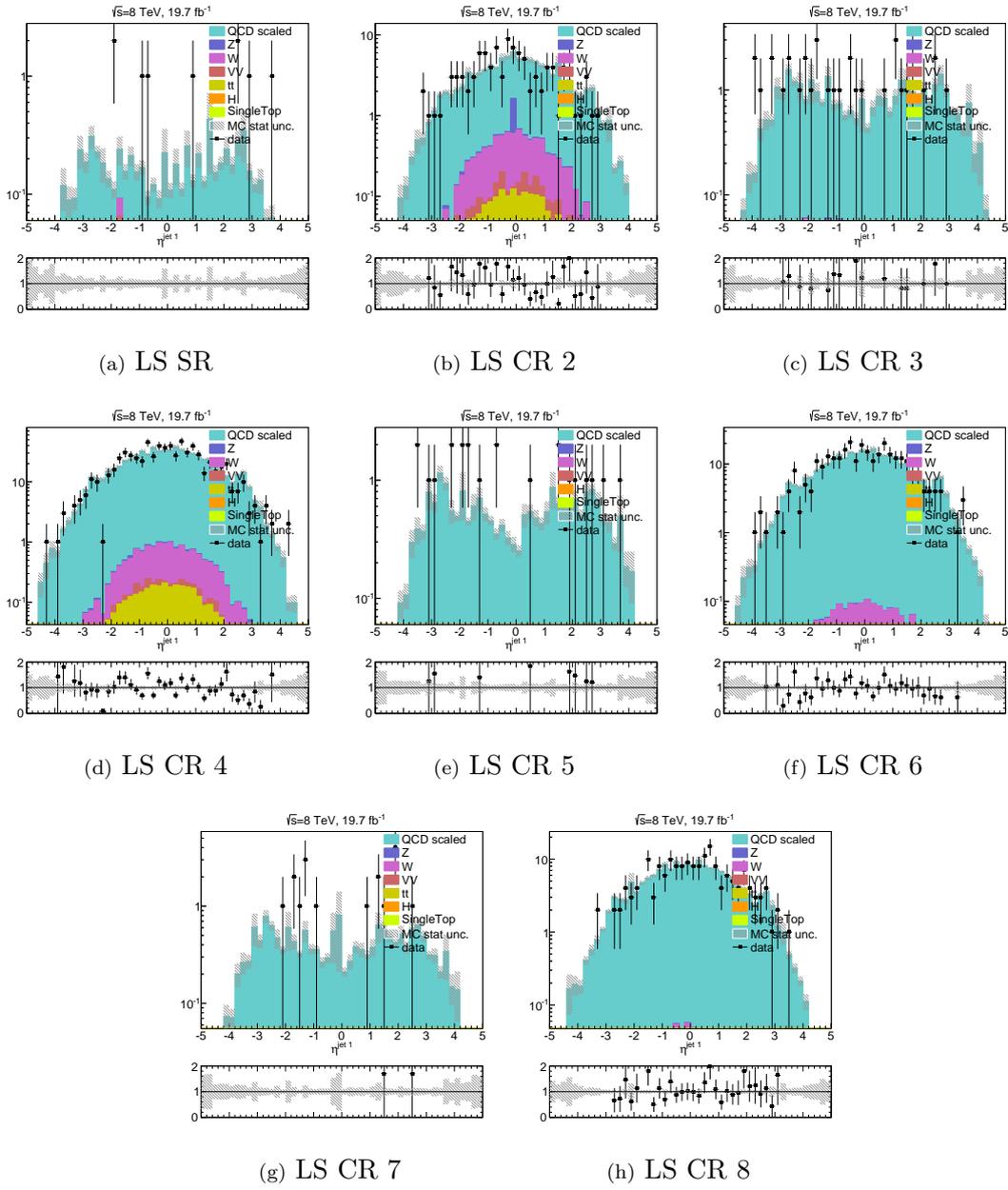
FIGURE B.54: Logarithmic pseudorapidity of the leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
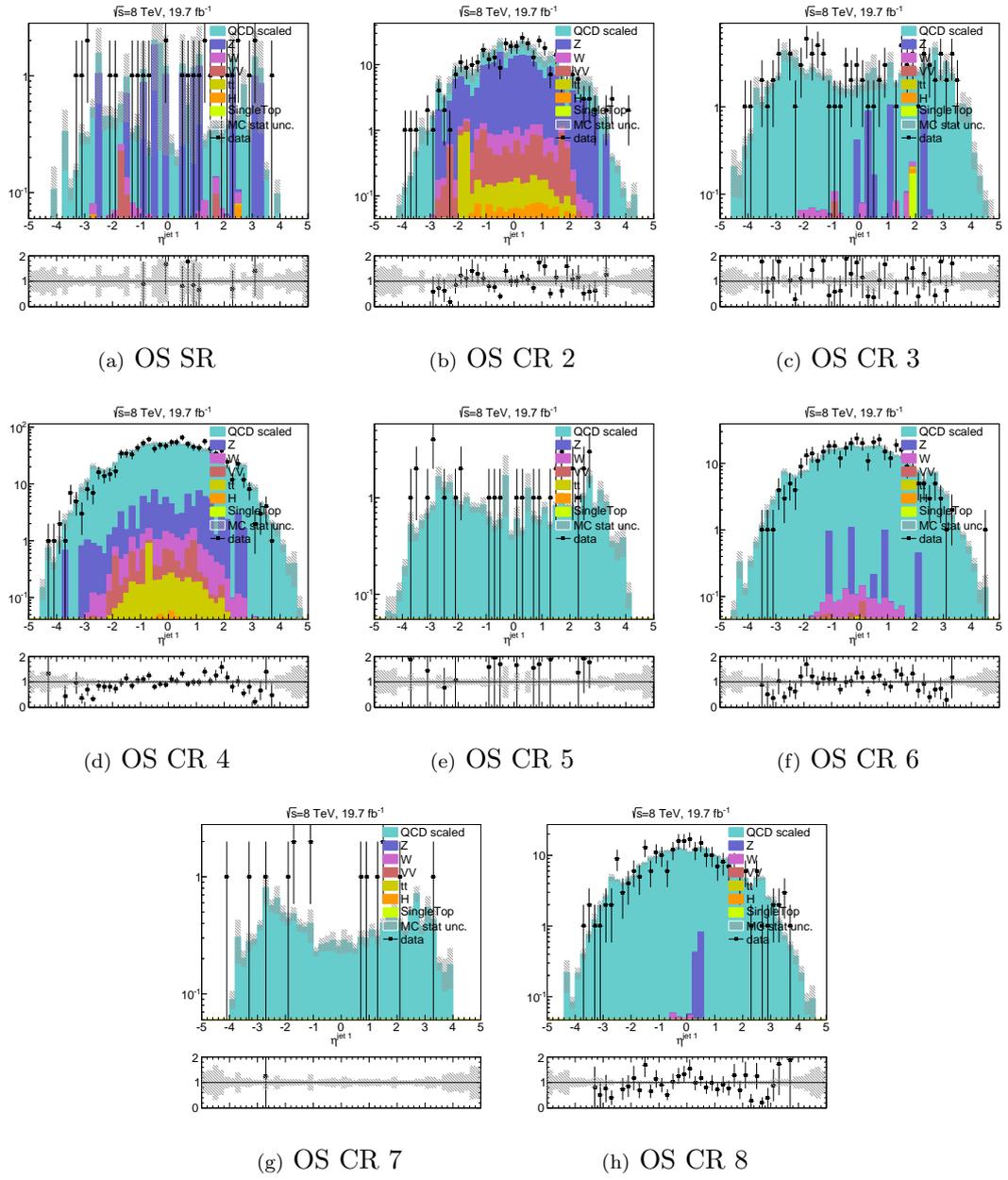
FIGURE B.55: Logarithmic pseudorapidity of the leading jet for OS

B.4.14 $\eta^{jet\ 2}$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
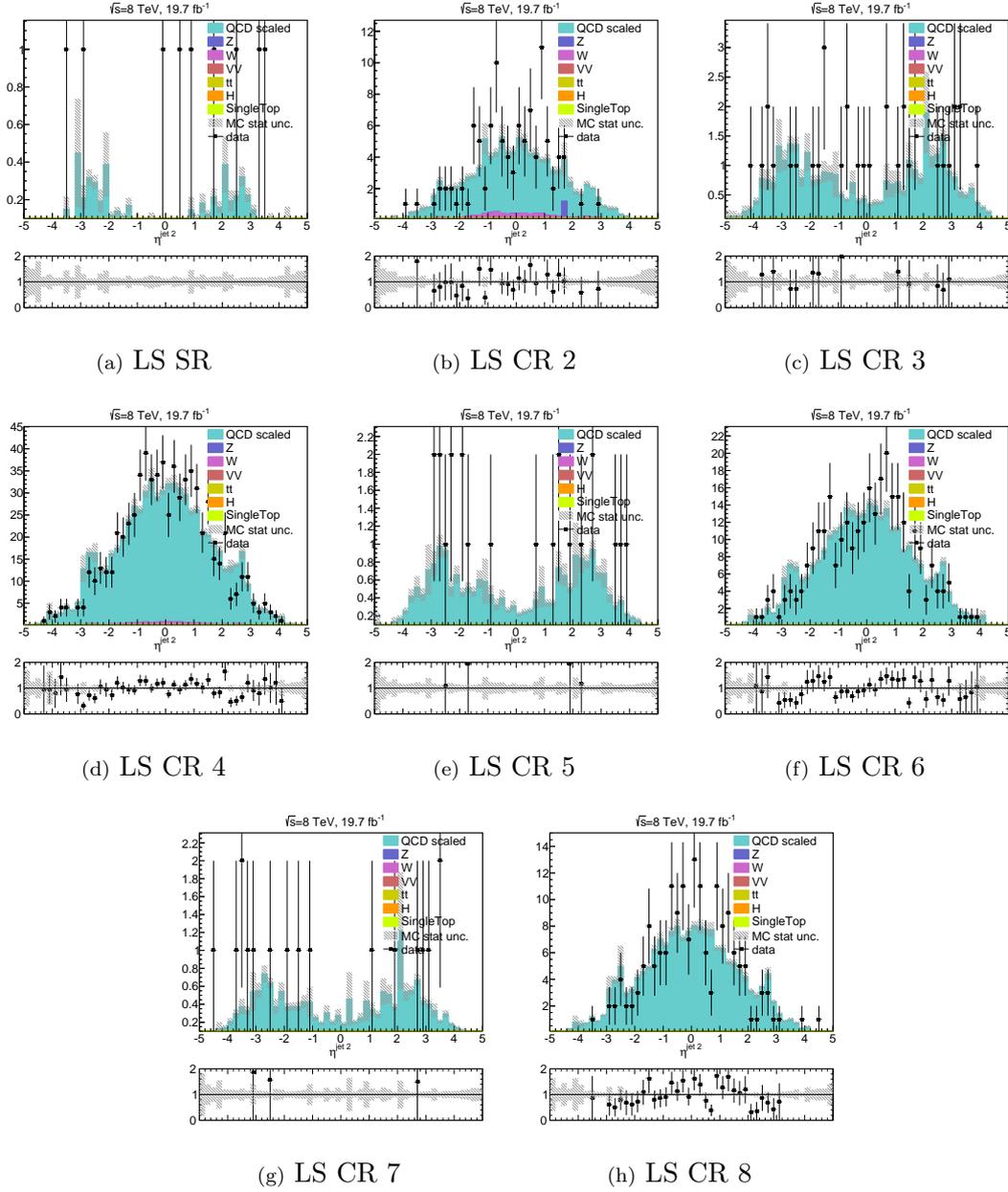
FIGURE B.56: Linear pseudorapidity of the second to leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
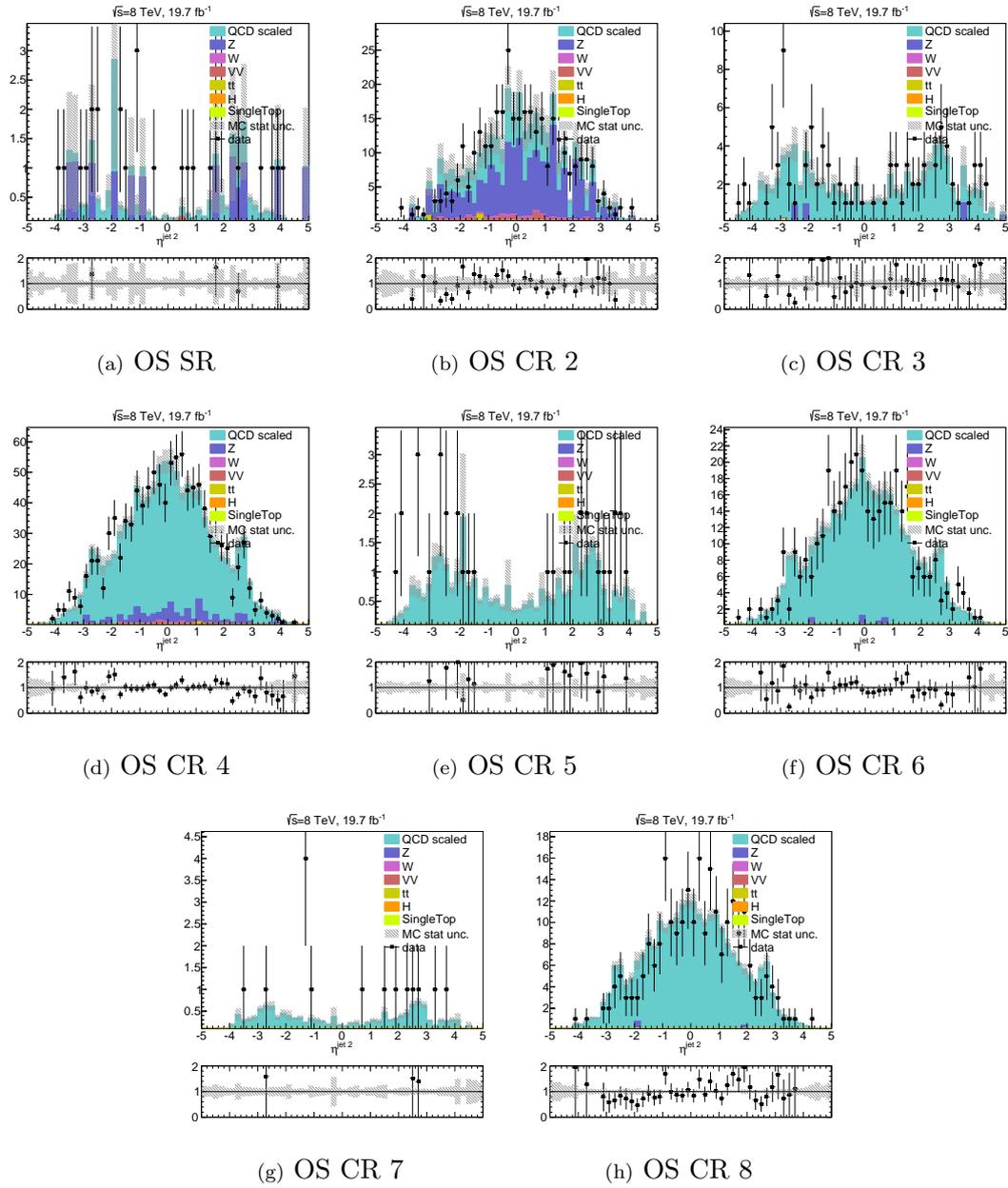
FIGURE B.57: Linear pseudorapidity of the second to leading jet for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
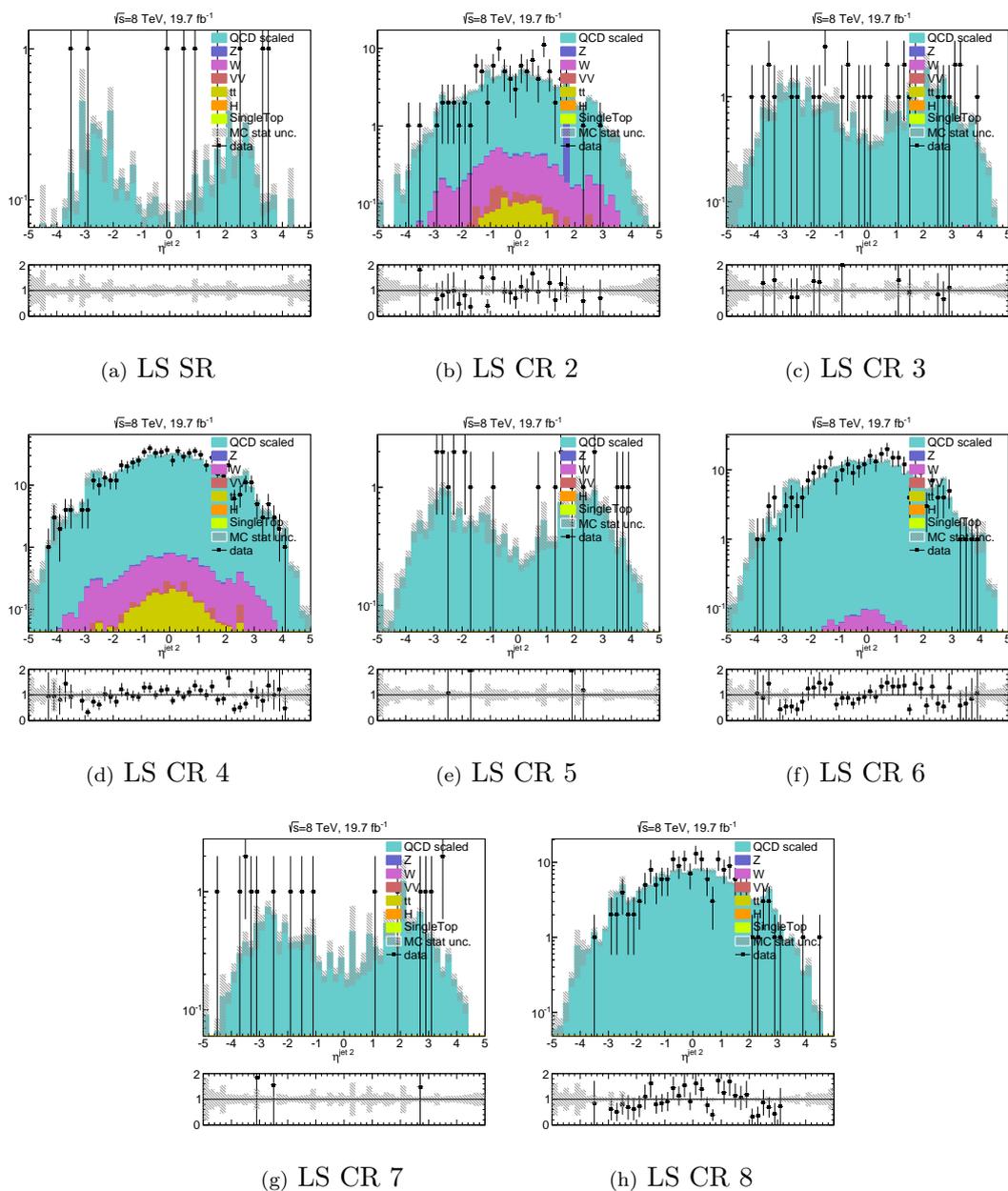
FIGURE B.58: Logarithmic pseudorapidity of the second to leading jet for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
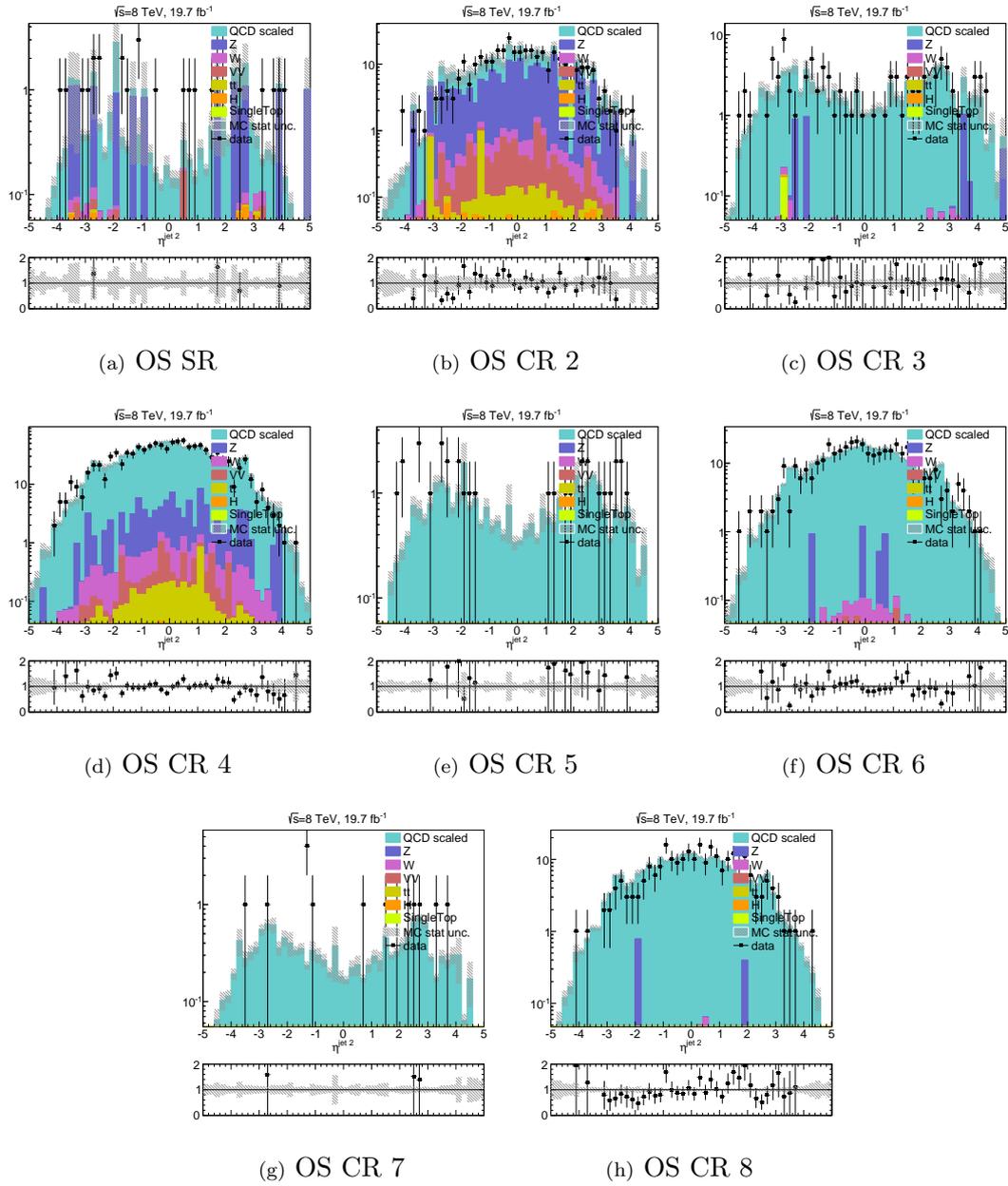
FIGURE B.59: Logarithmic pseudorapidity of the second to leading jet for OS

B.4.15   Inclusive η^{jet}

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
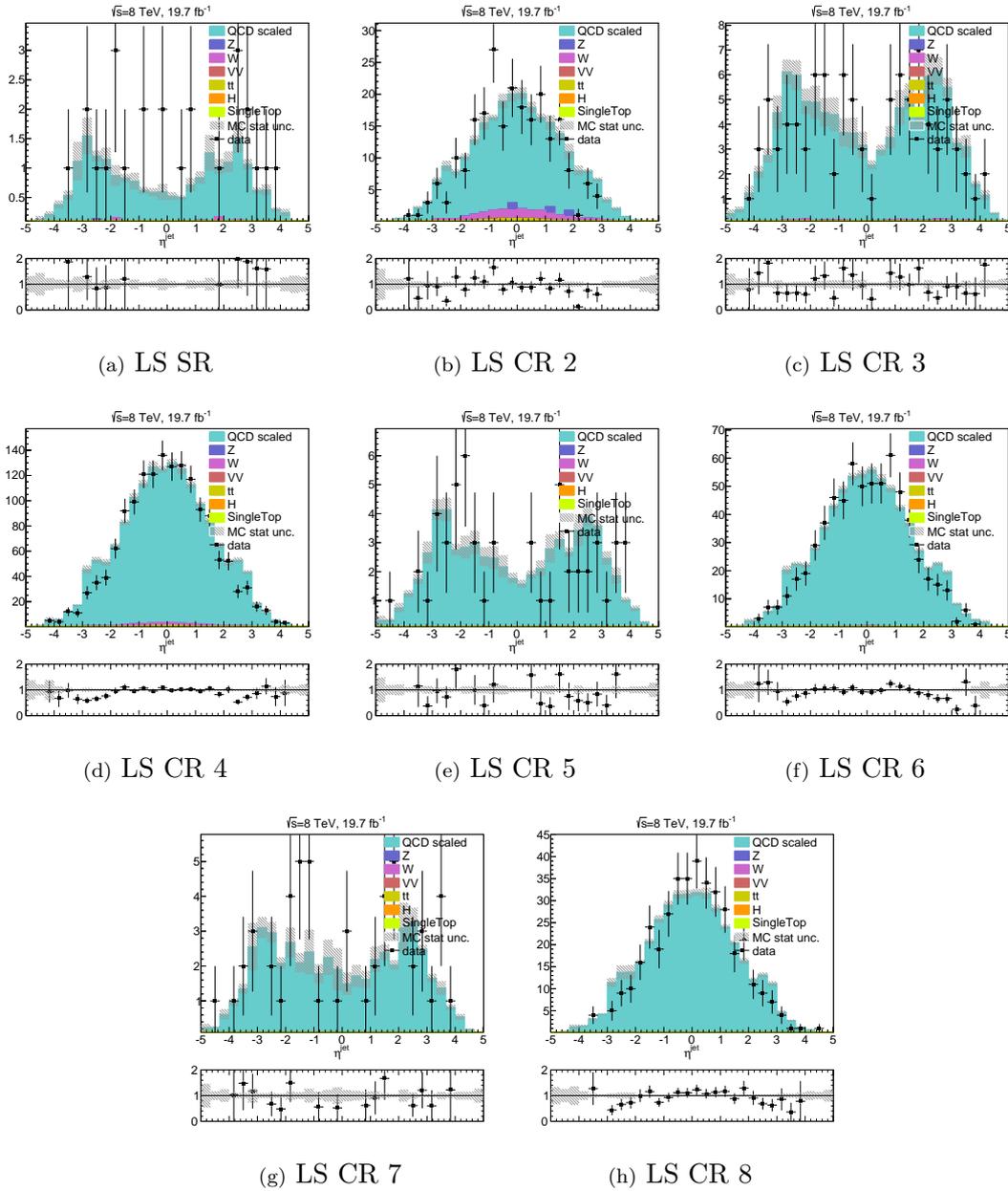
FIGURE B.60: Linear pseudorapidity of all jets for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
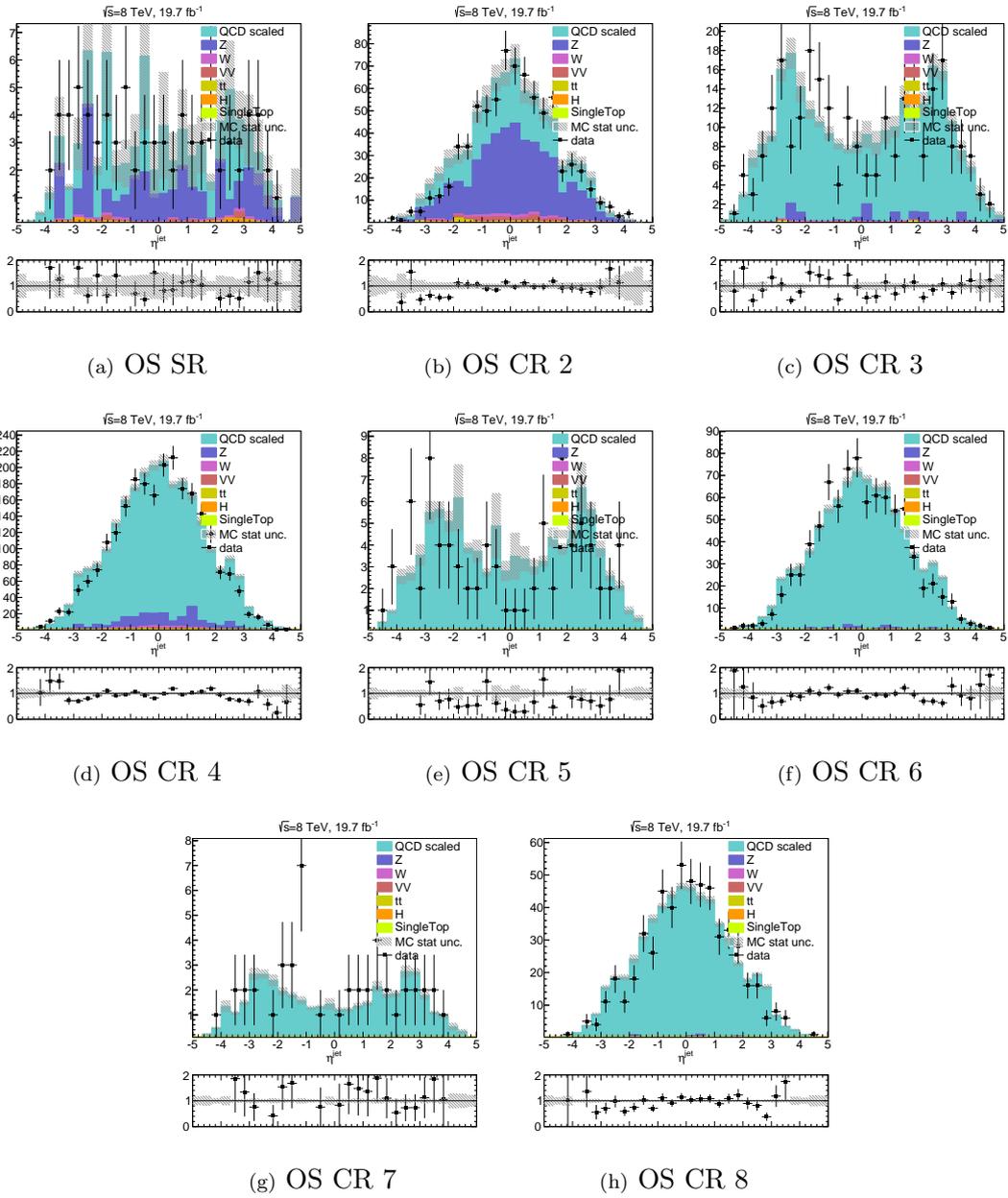
FIGURE B.61: Linear pseudorapidity of all jets for OS

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
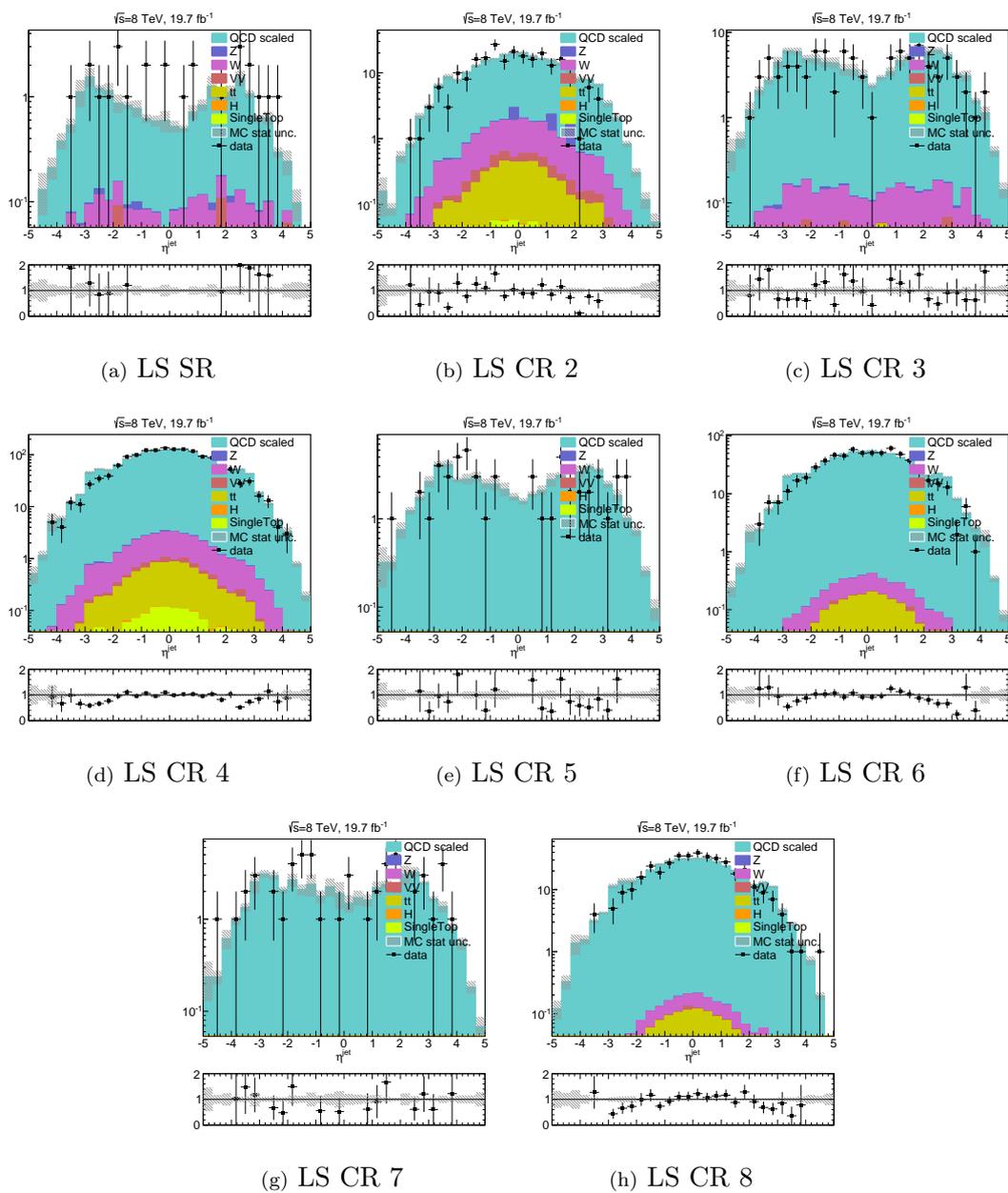
FIGURE B.62: Logarithmic pseudorapidity of all jets for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
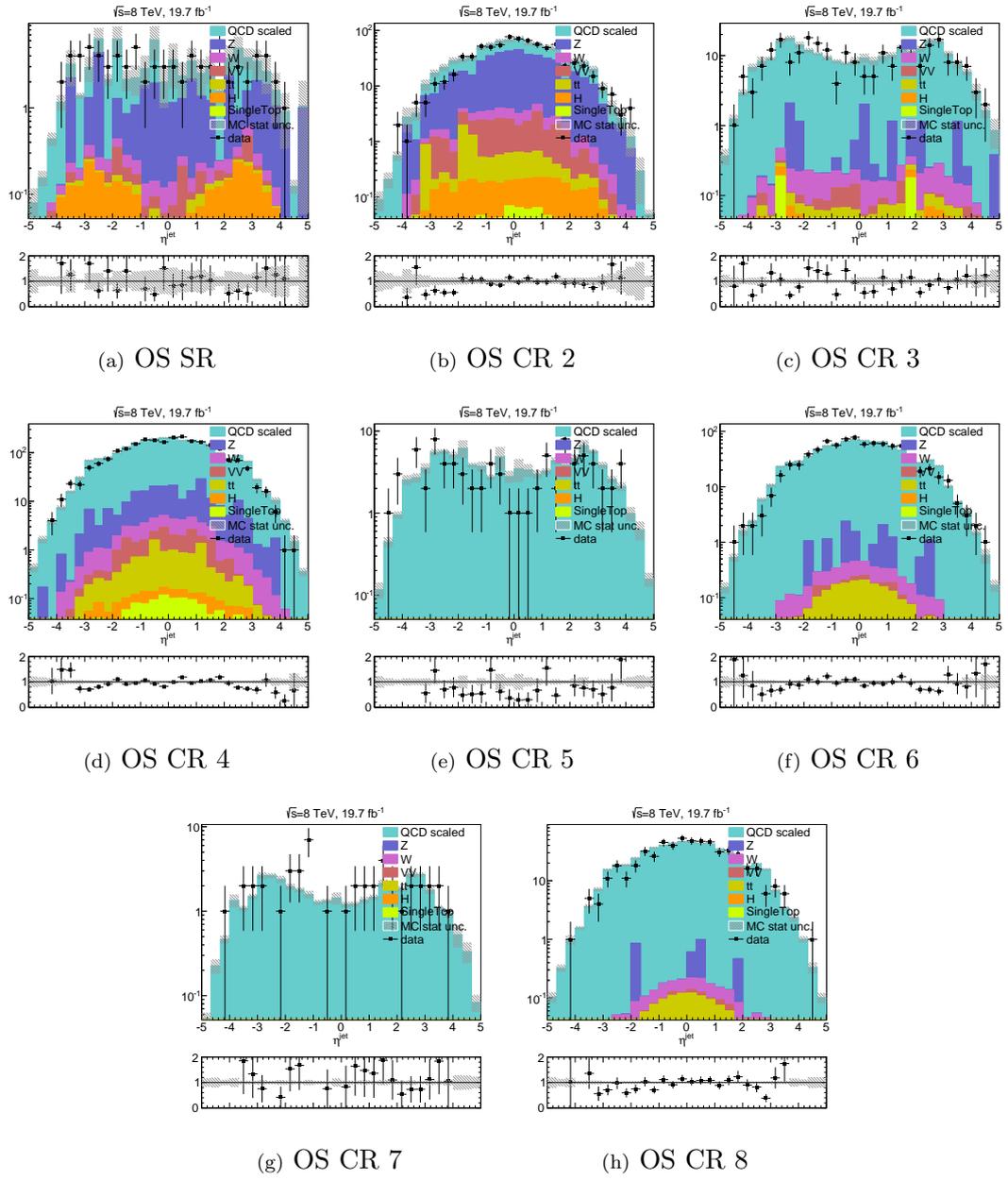
FIGURE B.63: Logarithmic pseudorapidity of all jets for OS

B.4.16 Missing transverse momentum

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
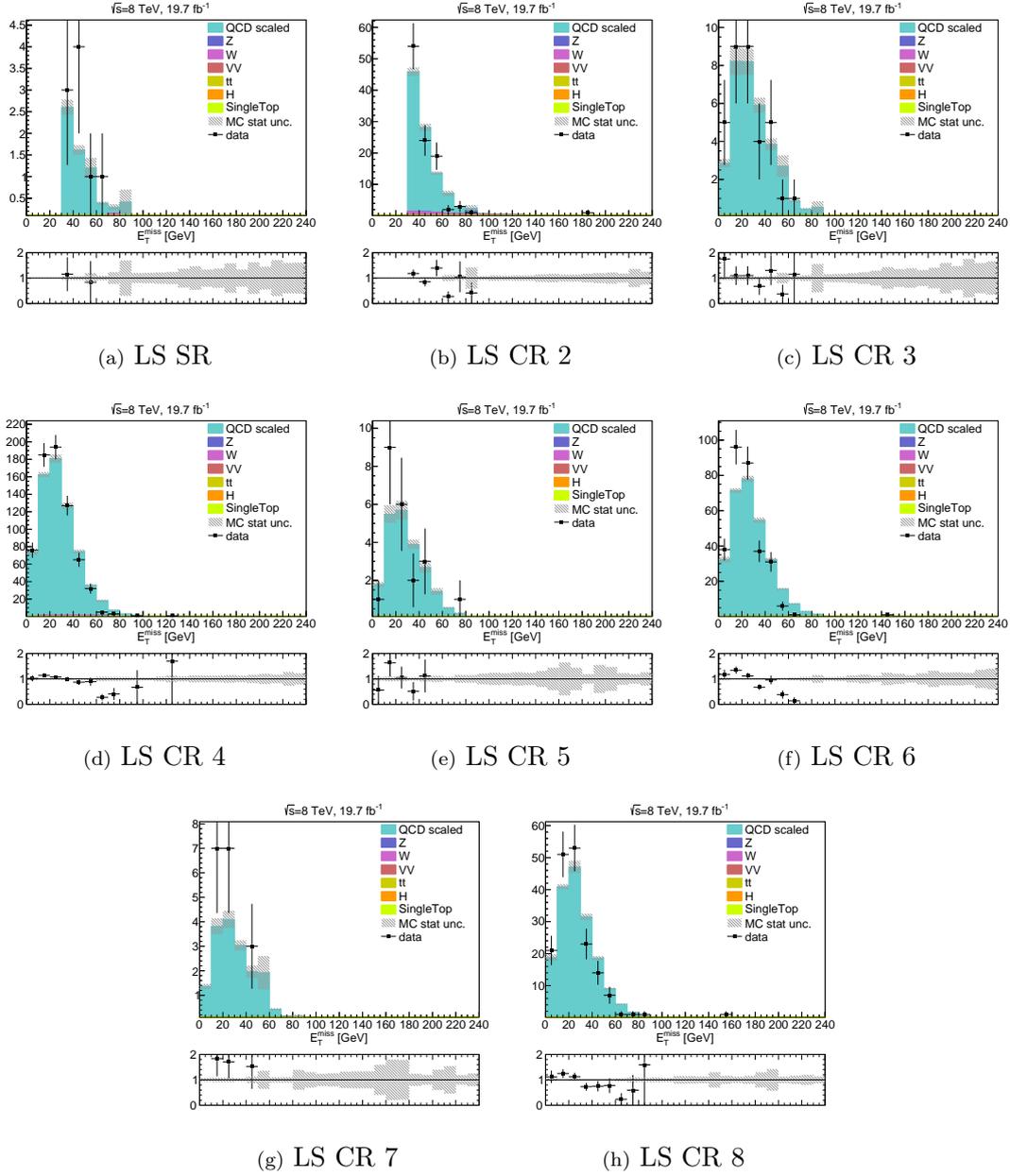
FIGURE B.64: Linear vectorial sum of reconstructed object transverse momenta for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
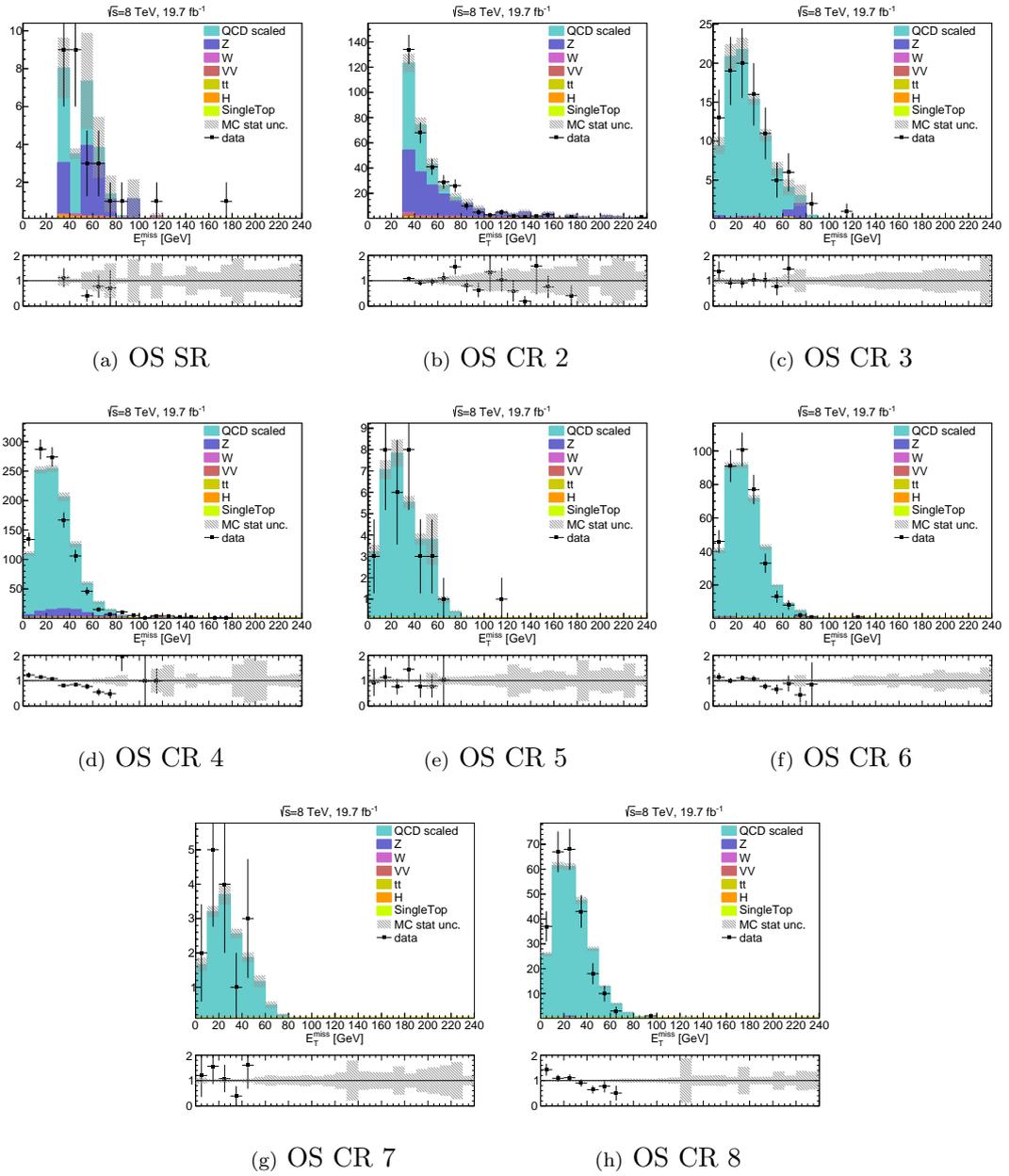
FIGURE B.65: Linear vectorial sum of reconstructed object transverse momenta for OS

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
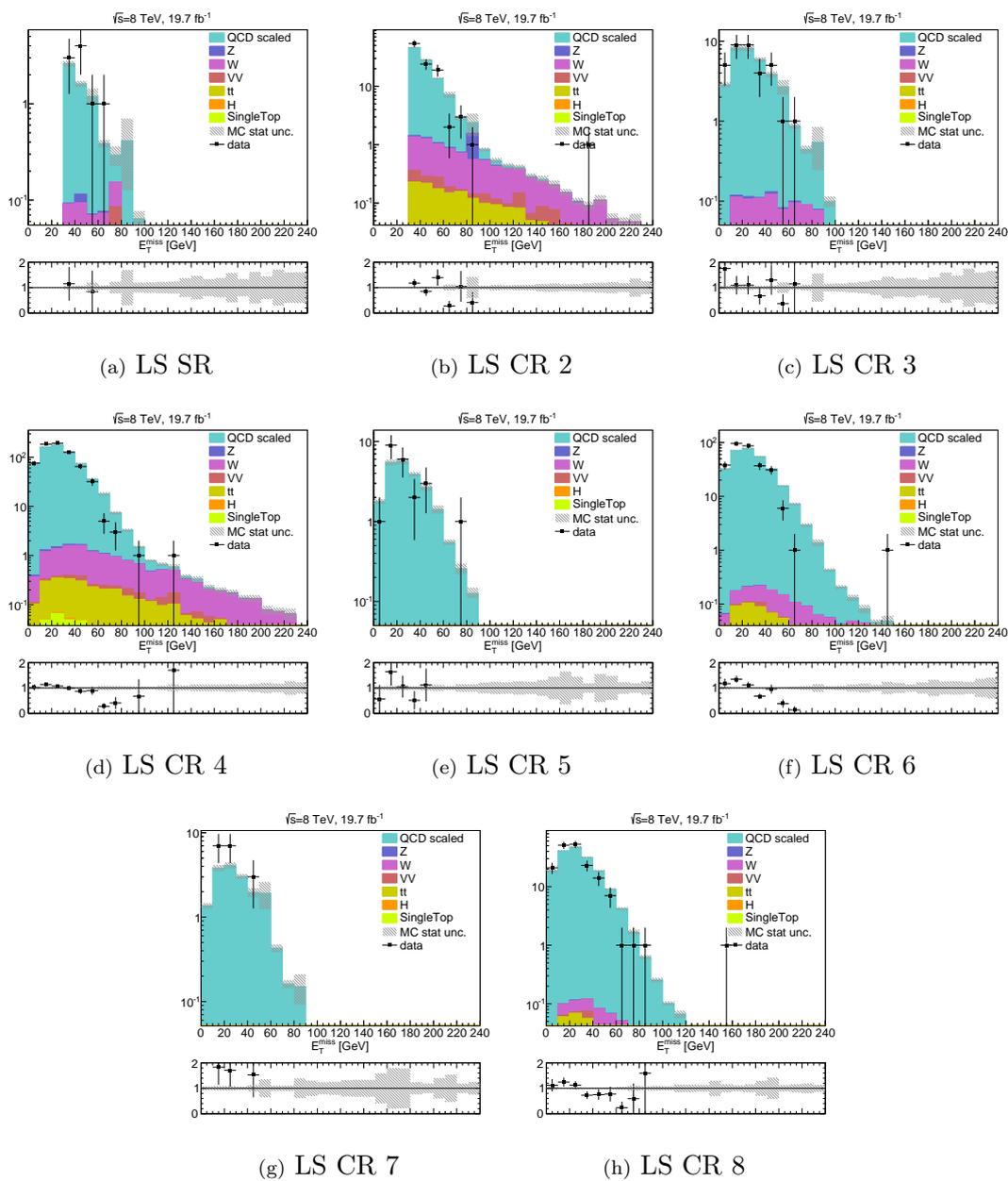
FIGURE B.66: Logarithmic vectorial sum of reconstructed object transverse momenta for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6
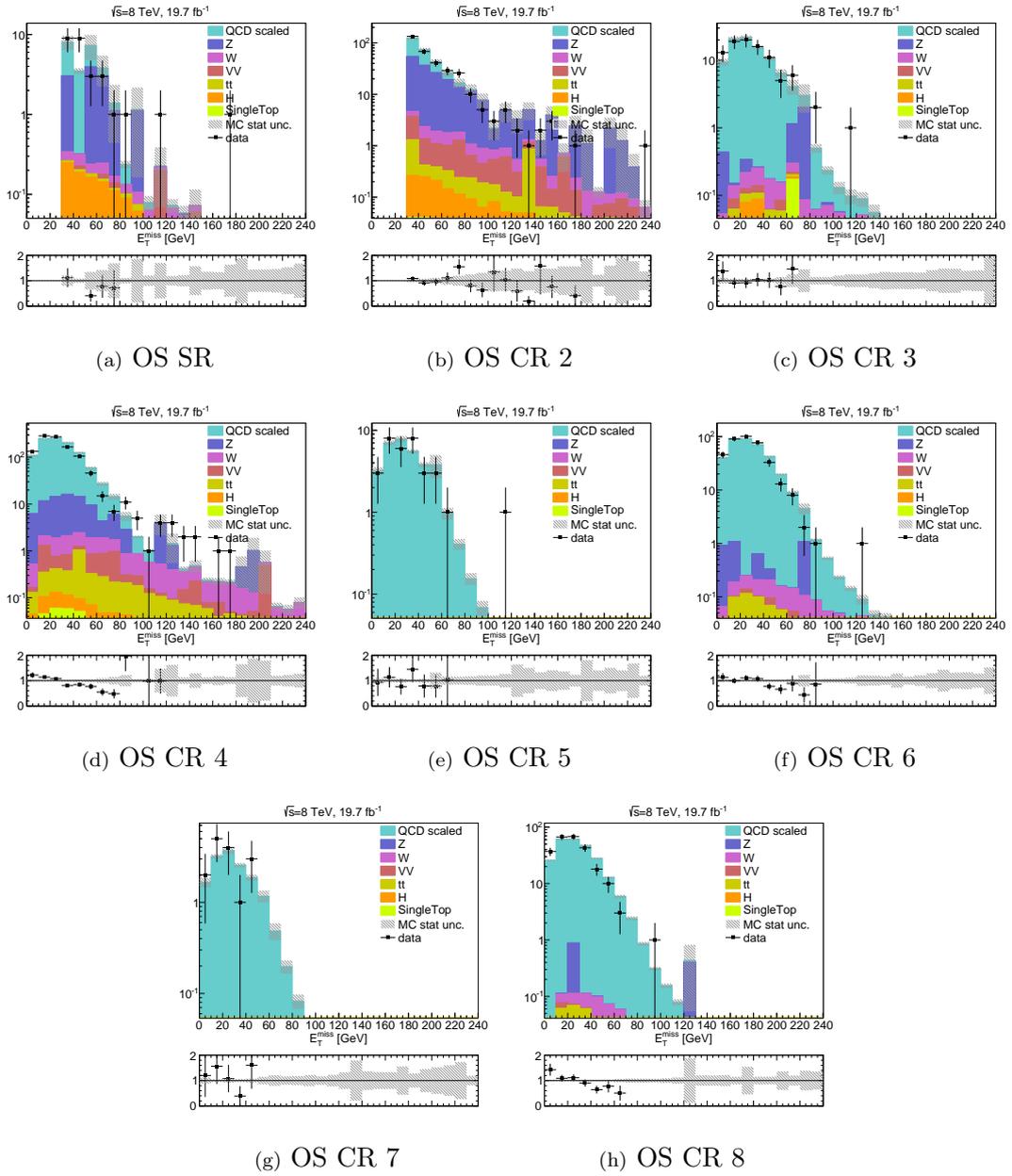
(g) OS CR 7

(h) OS CR 8

FIGURE B.67: Logarithmic vectorial sum of reconstructed object transverse momenta for OS

B.4.17 Number of jets

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
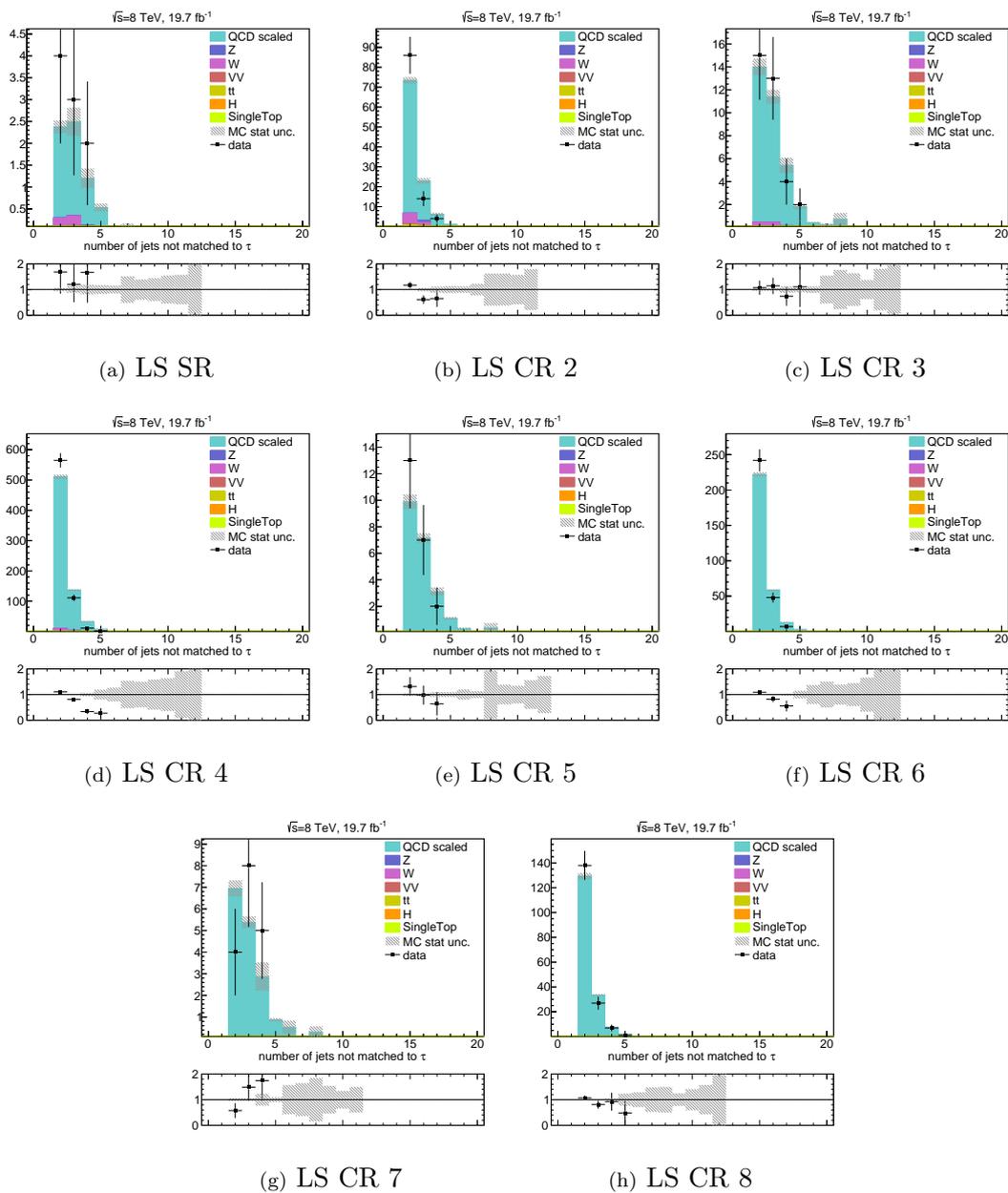
FIGURE B.68: Linear number of jets for LS
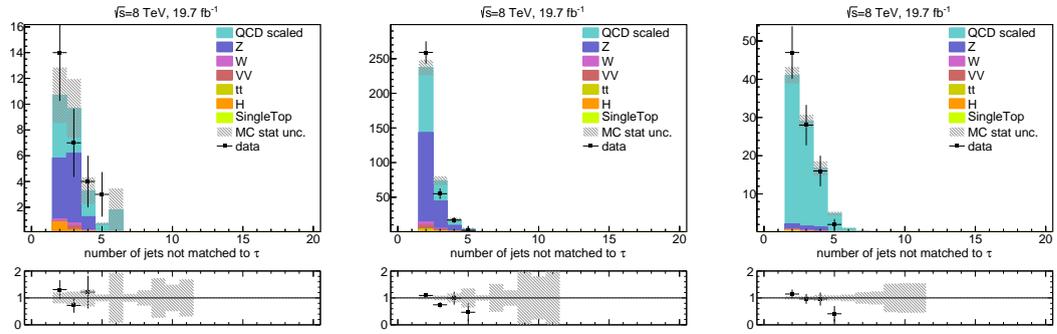
(a) OS SR

(b) OS CR 2

(c) OS CR 3
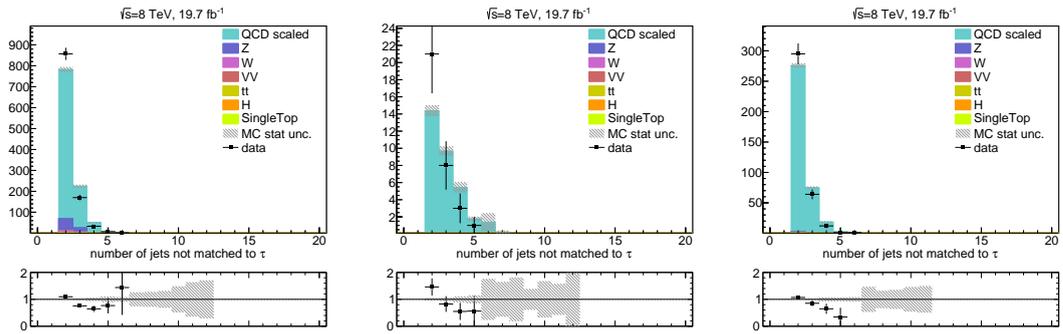
(d) OS CR 4

(e) OS CR 5

(f) OS CR 6
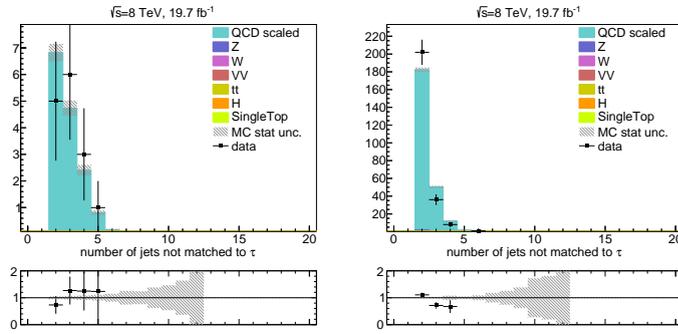
(g) OS CR 7

(h) OS CR 8

FIGURE B.69: Linear number of jets for OS

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
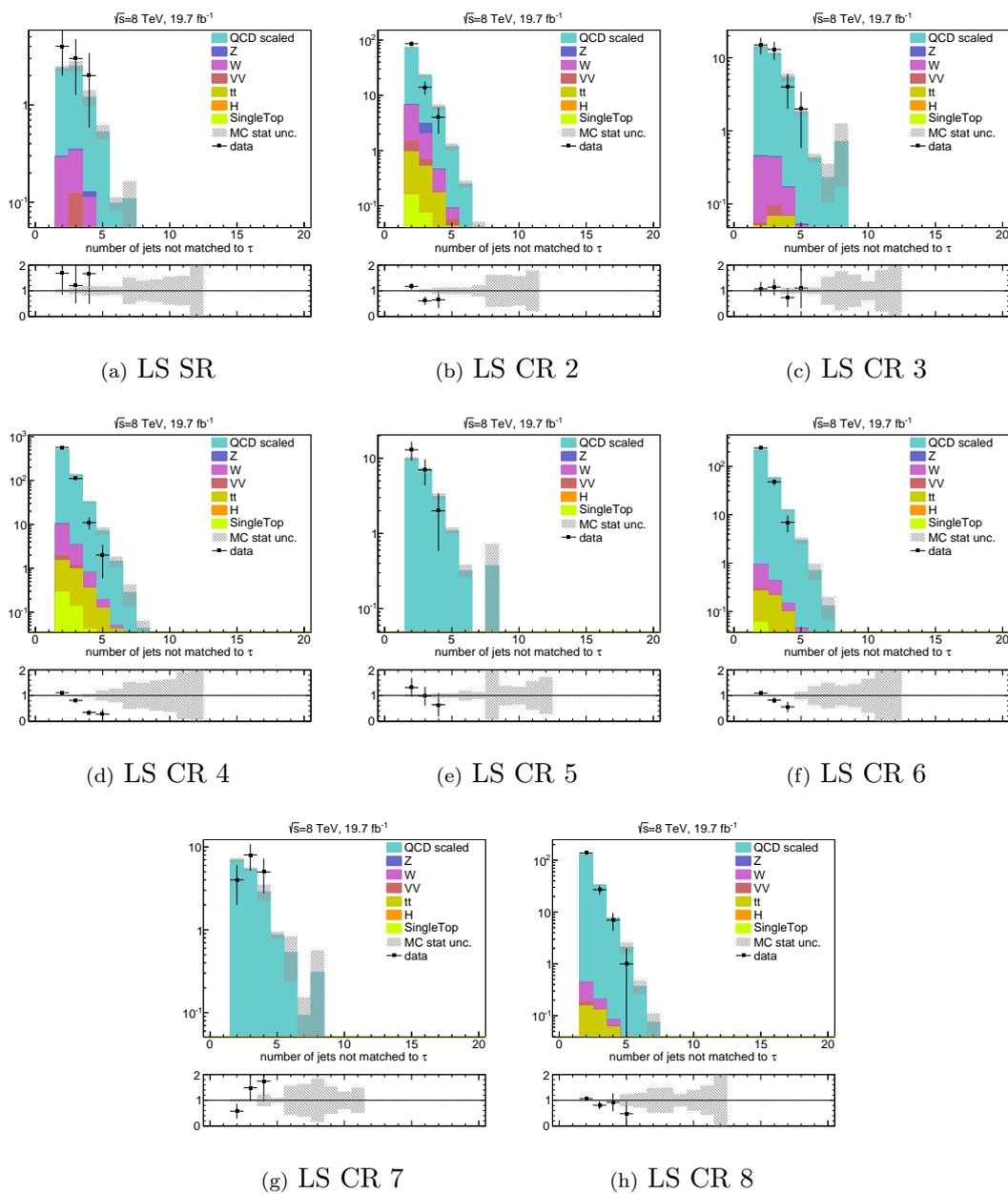
FIGURE B.70: Logarithmic number of jets for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
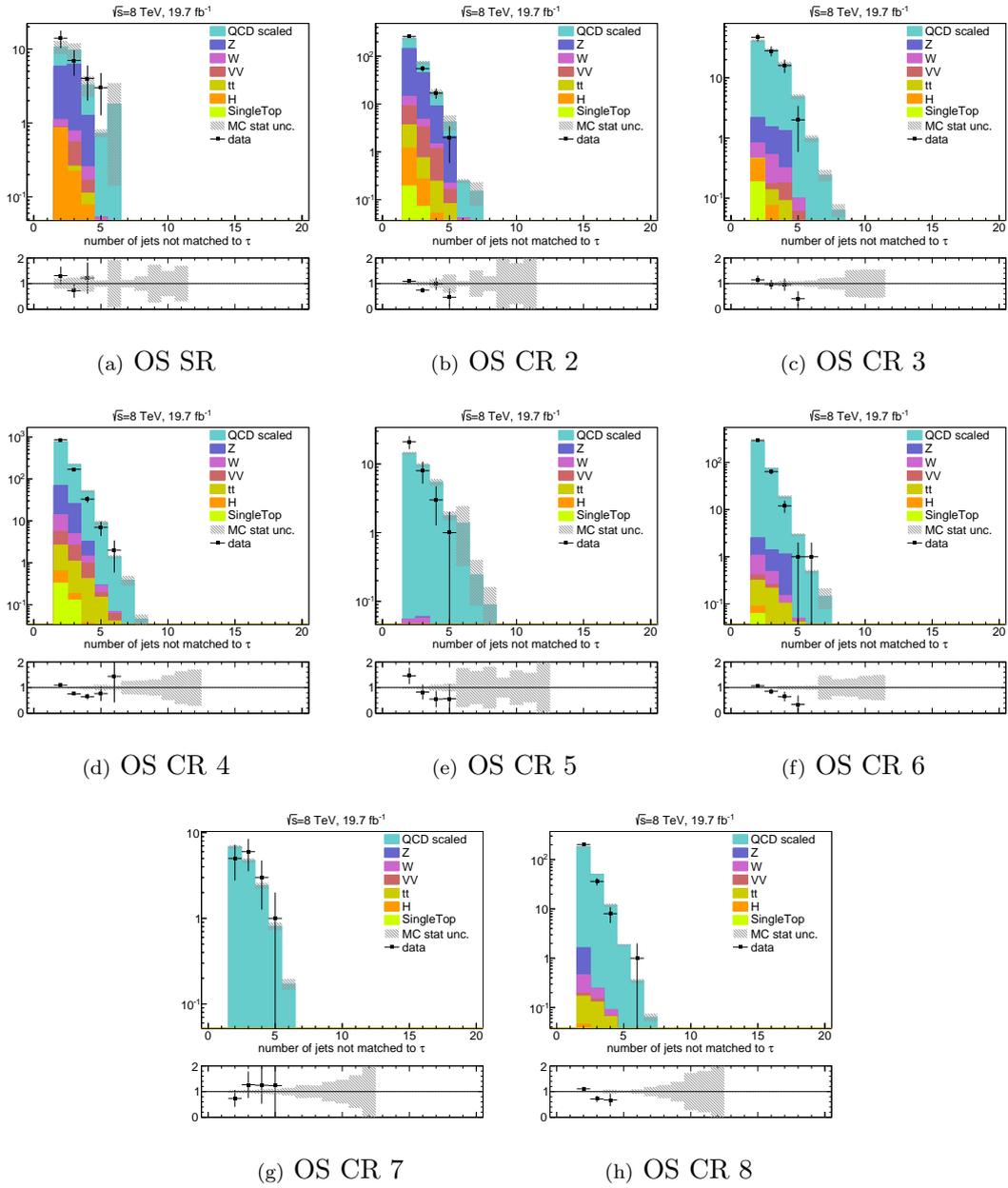
FIGURE B.71: Logarithmic number of jets for OS

B.4.18   Number of reconstructed Vertices

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
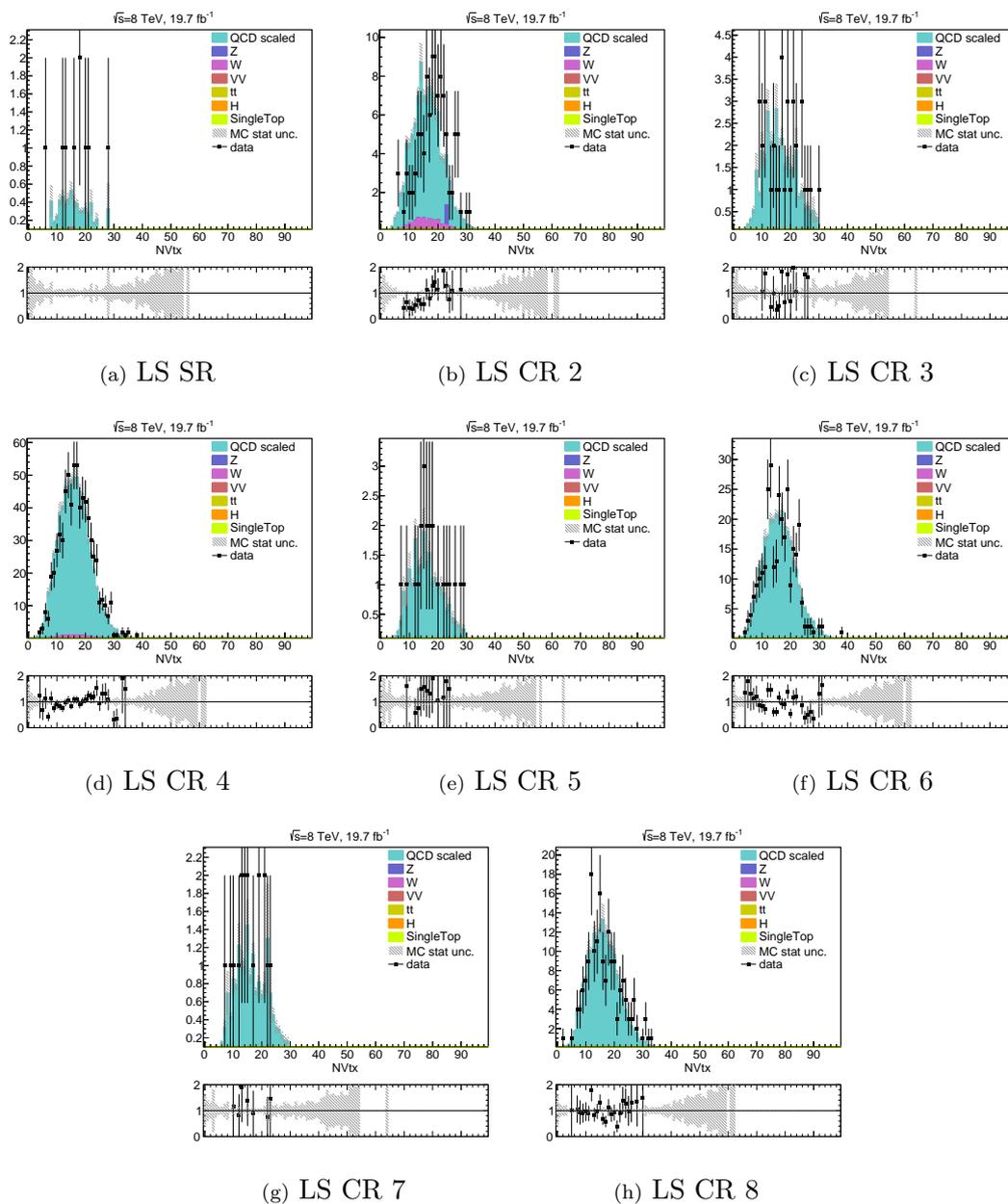
FIGURE B.72: Linear number of reconstructed vertices for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
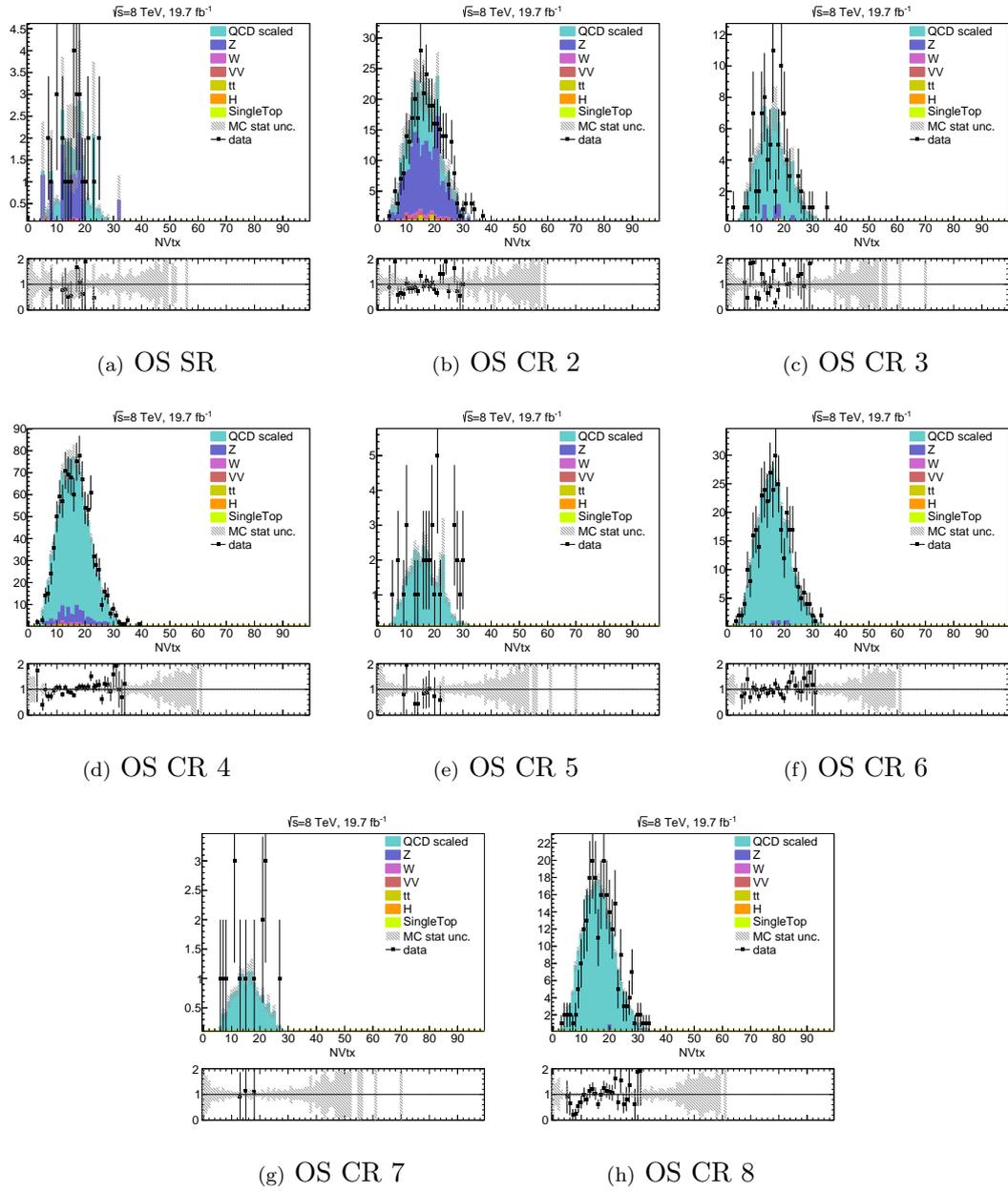
FIGURE B.73: Linear number of reconstructed vertices for OS

(a) LS SR   (b) LS CR 2   (c) LS CR 3

(d) LS CR 4   (e) LS CR 5   (f) LS CR 6

(g) LS CR 7   (h) LS CR 8
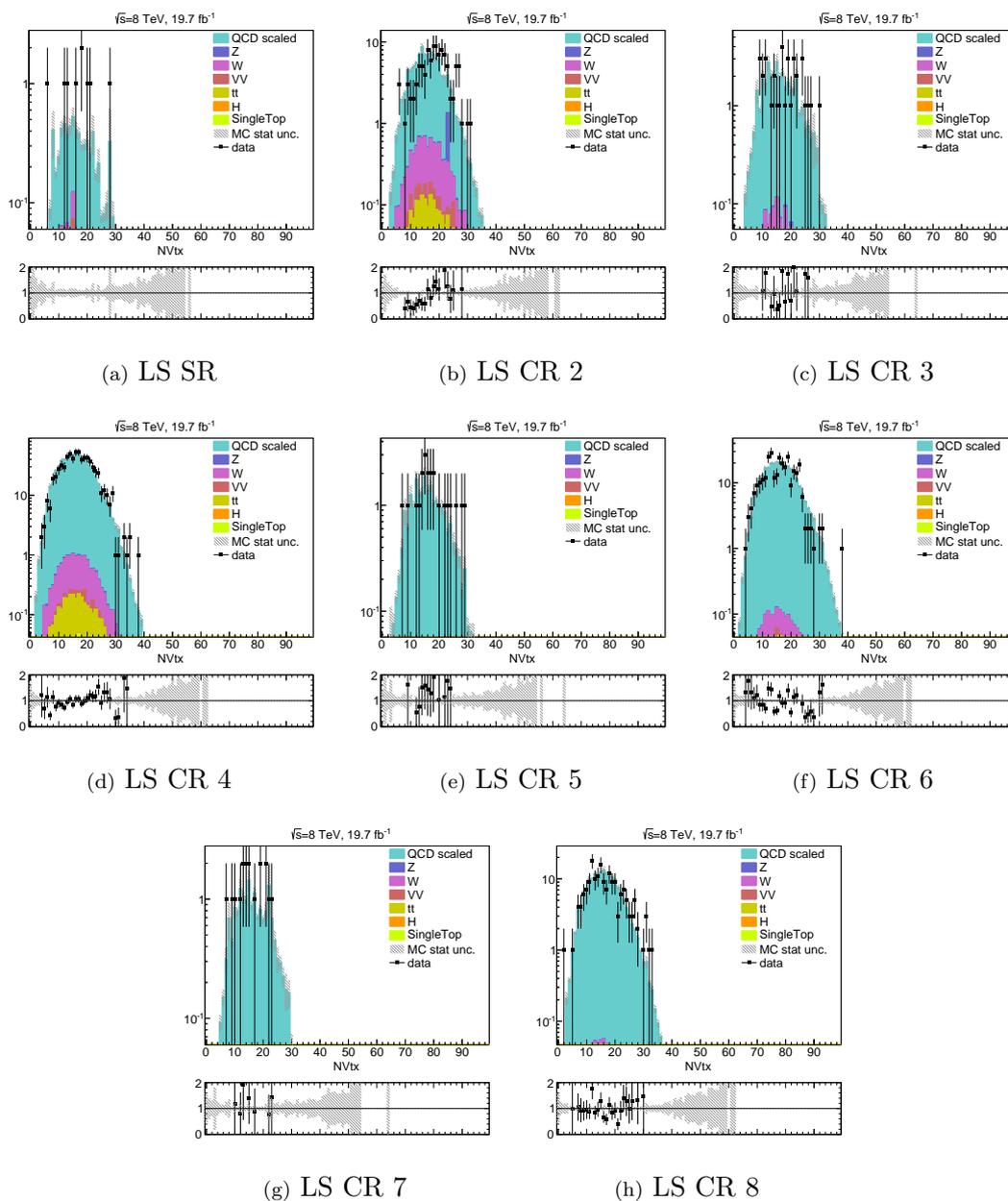
FIGURE B.74: Logarithmic number of reconstructed vertices for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
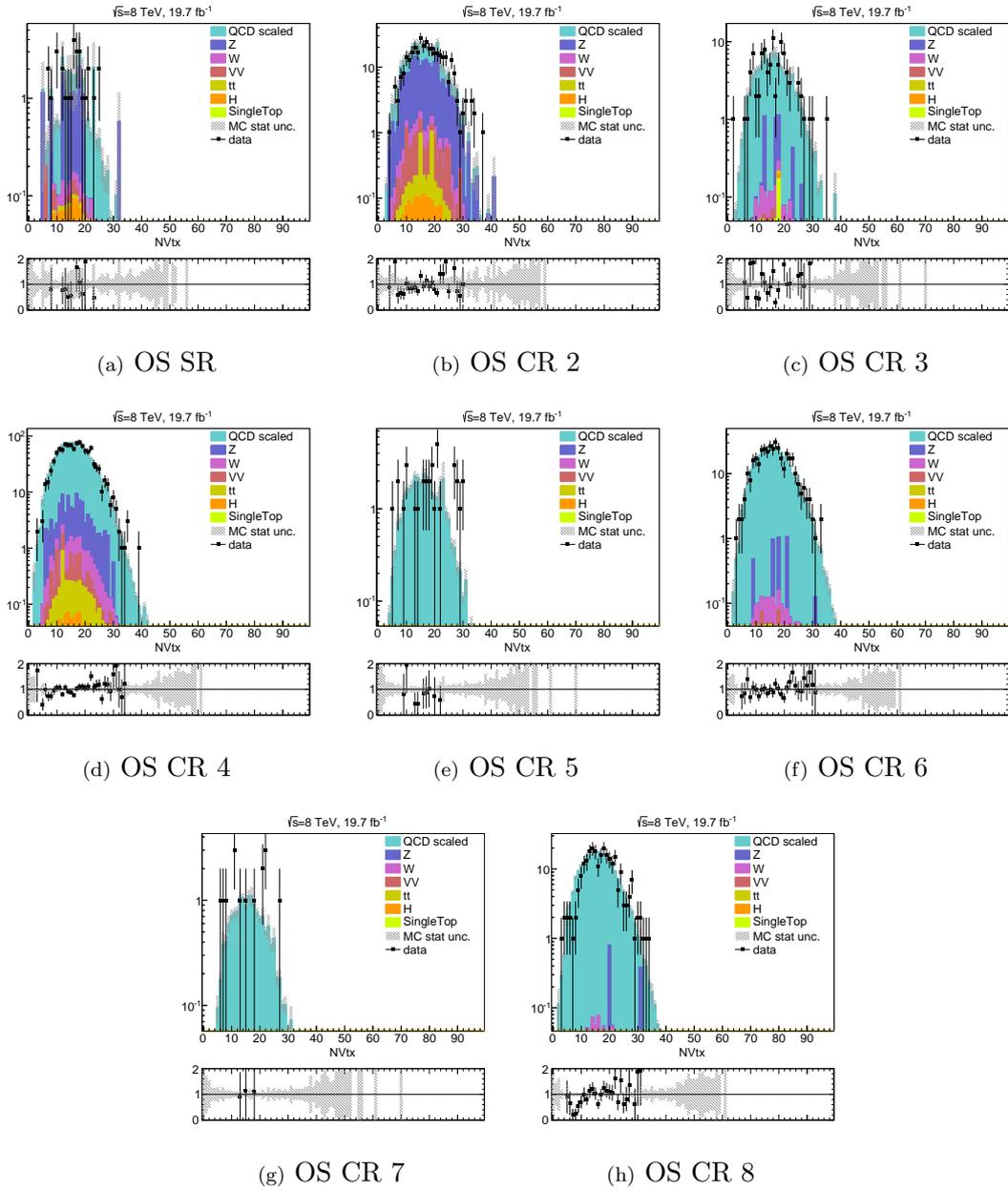
FIGURE B.75: Logarithmic number of reconstructed vertices for OS

B.4.19 p_T^τ 1

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
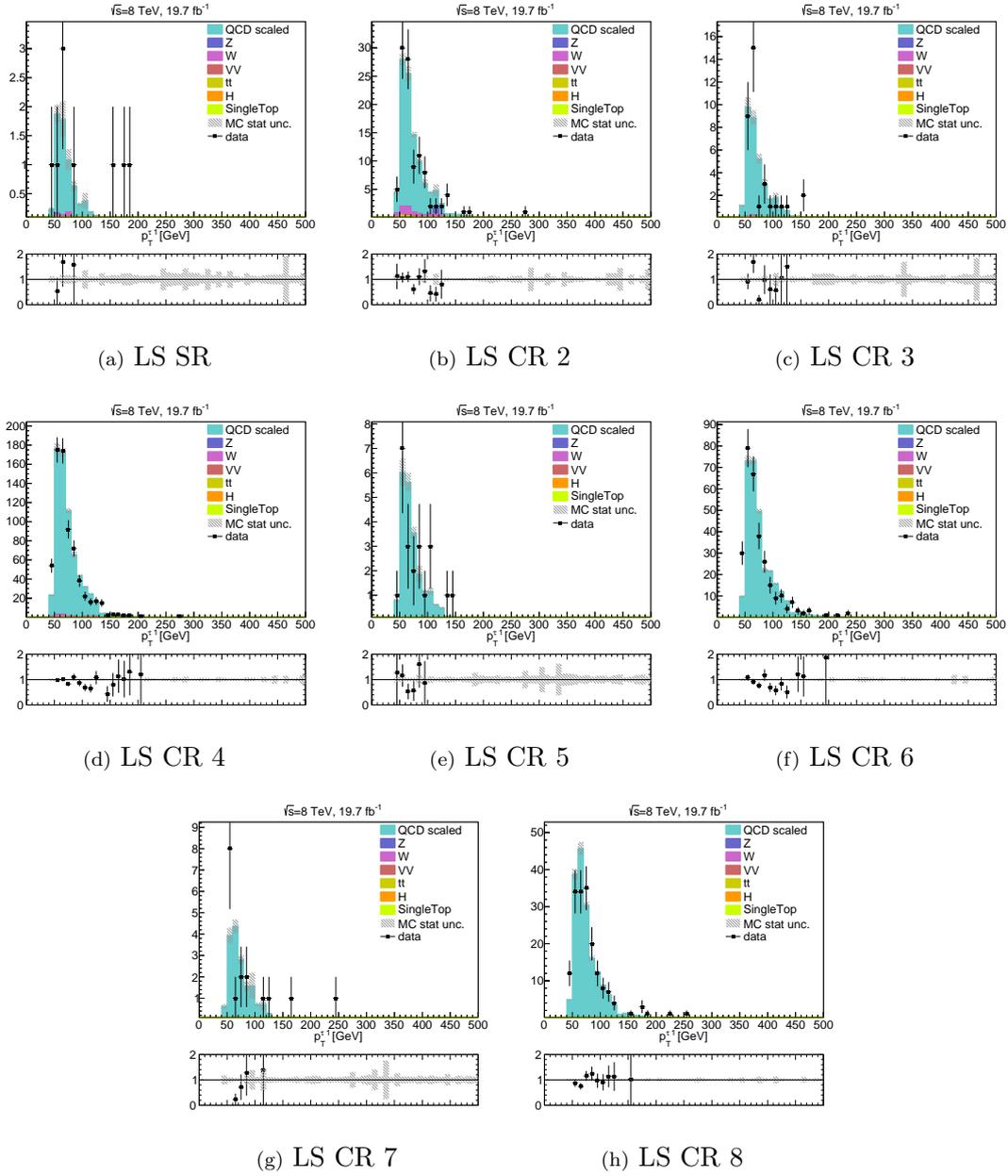
FIGURE B.76: Linear leading τ transverse momentum for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
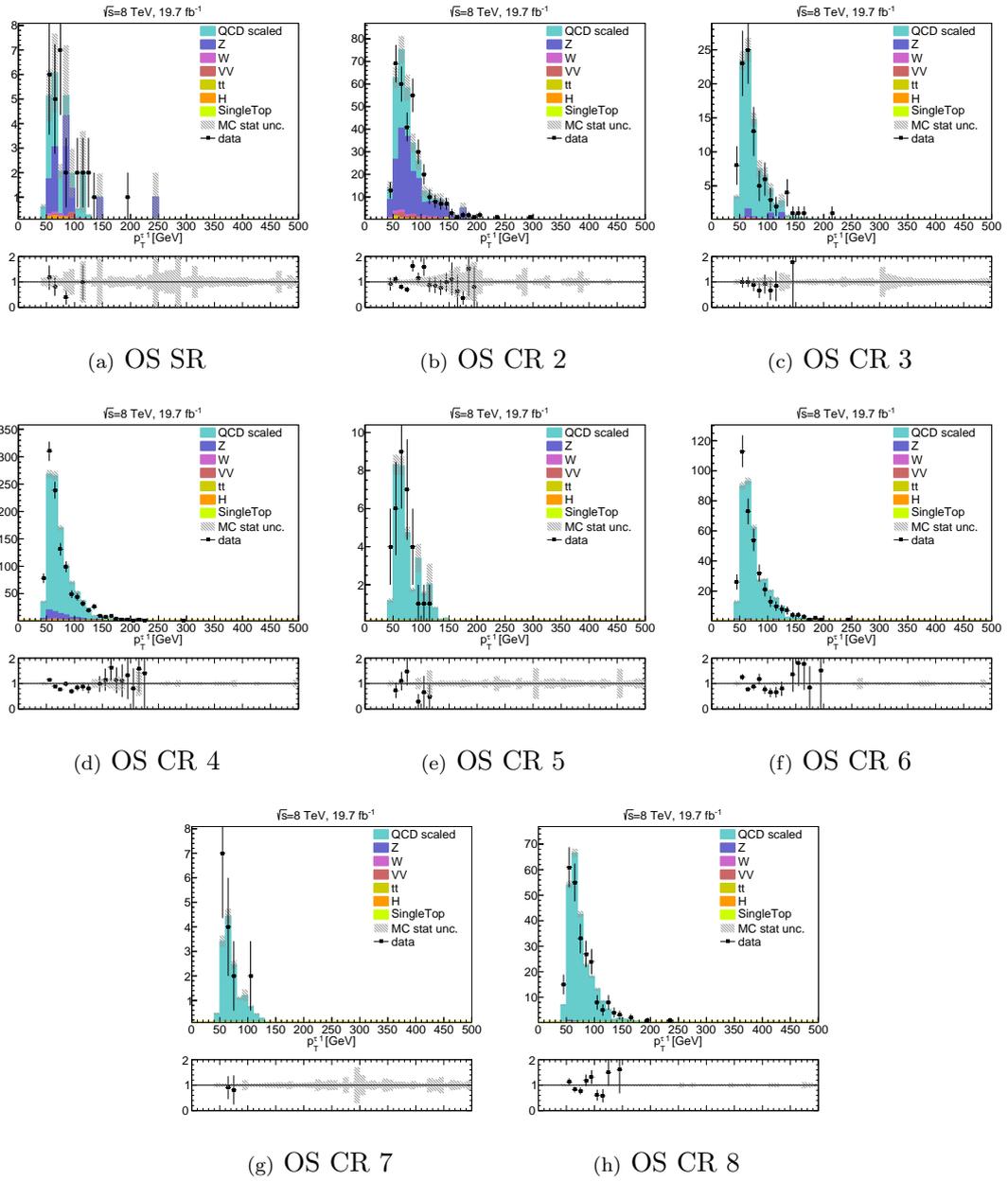
FIGURE B.77: Linear leading τ transverse momentum for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
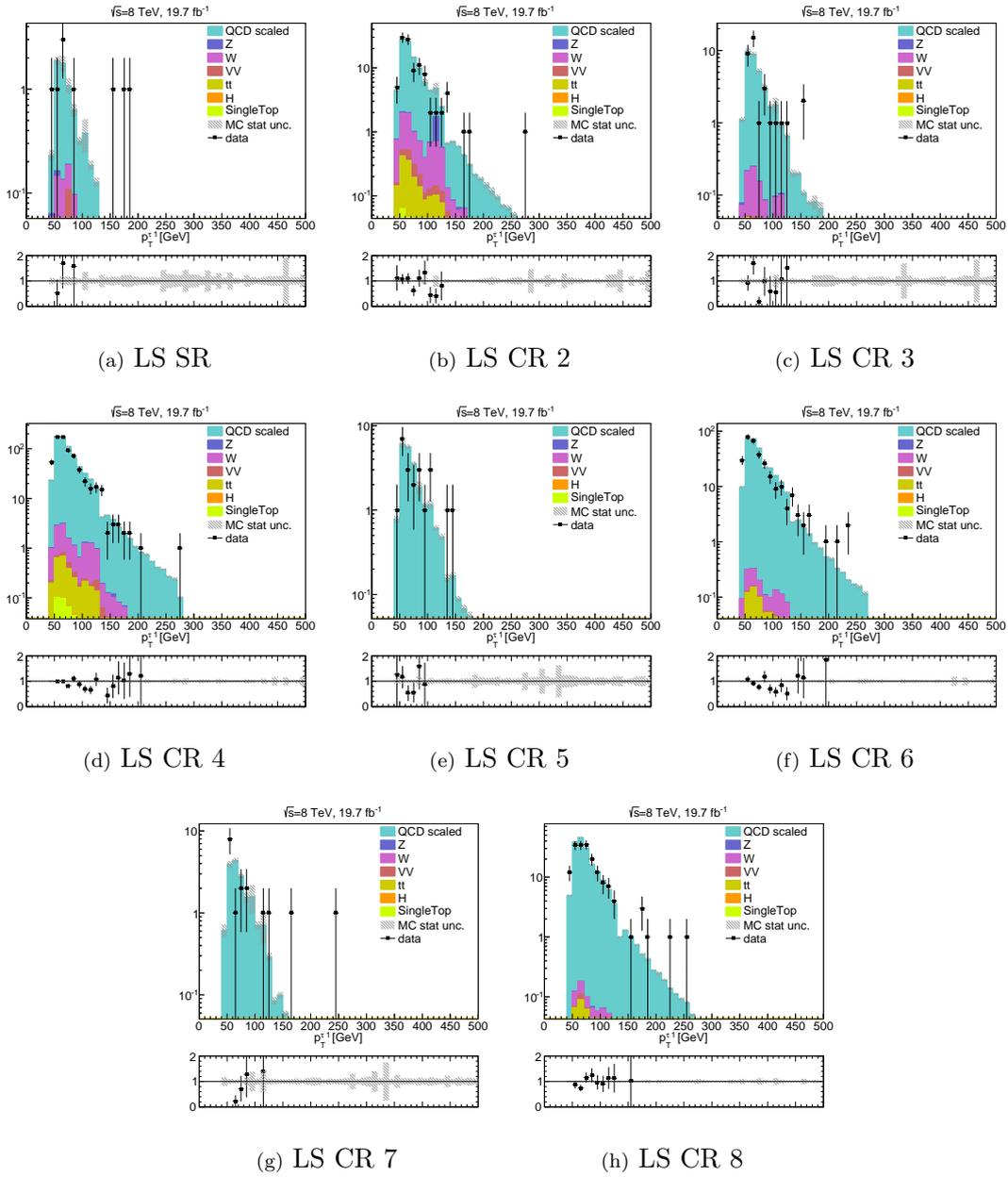
FIGURE B.78: Logarithmic leading τ transverse momentum for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
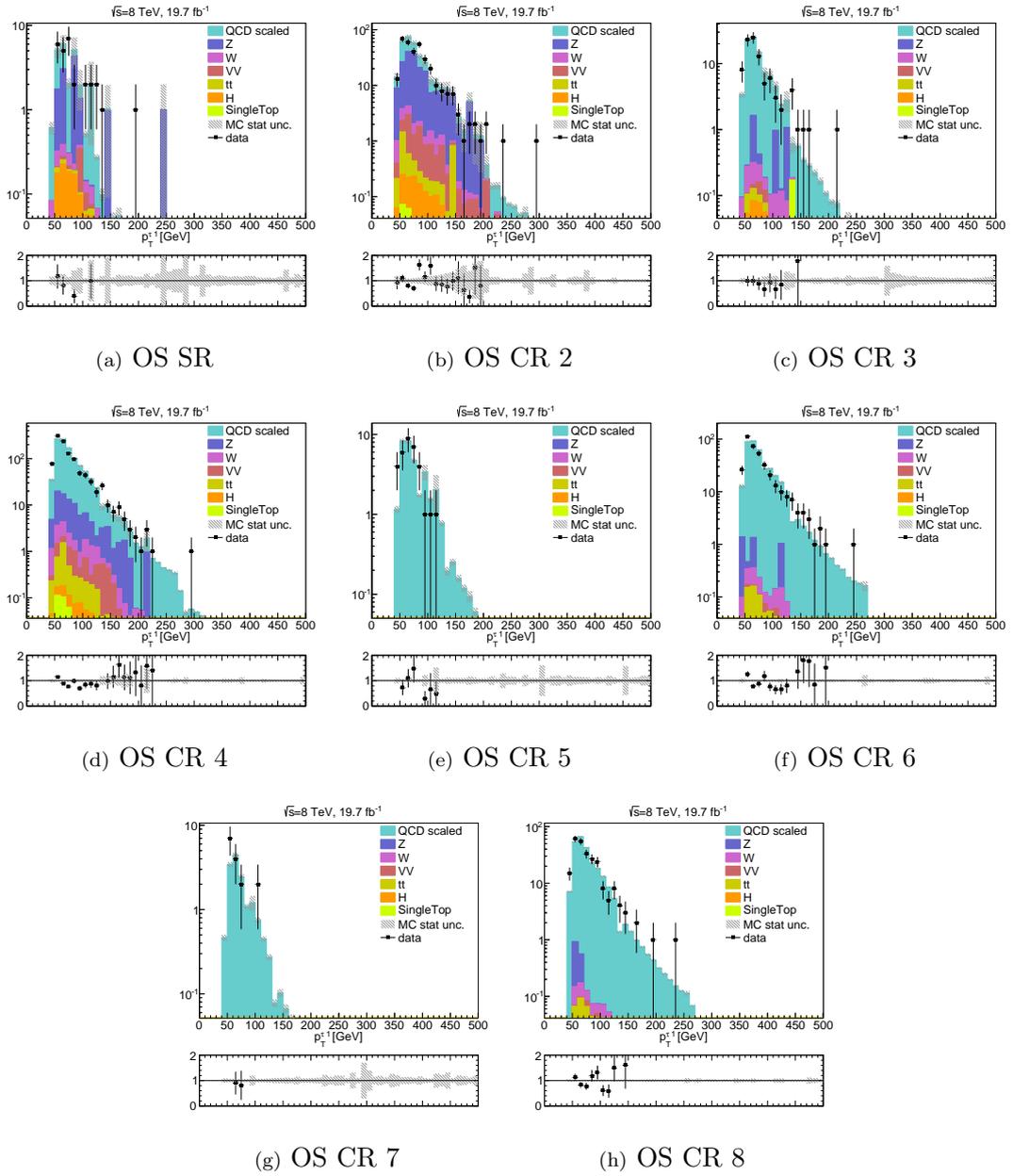
FIGURE B.79: Logarithmic leading τ transverse momentum for OS

B.4.20 p_T^τ 2

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
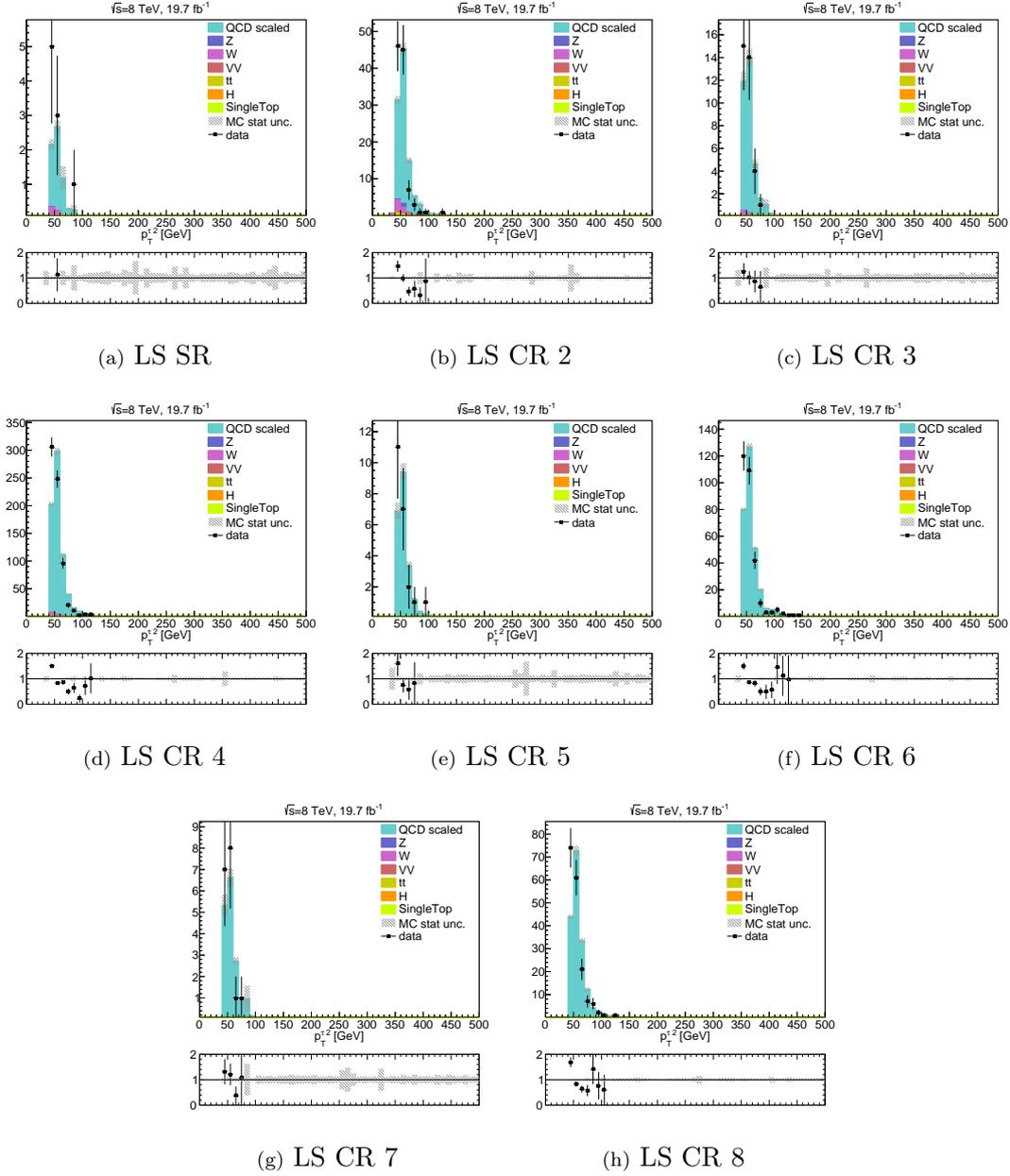
FIGURE B.80: Linear second to leading τ transverse momentum for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
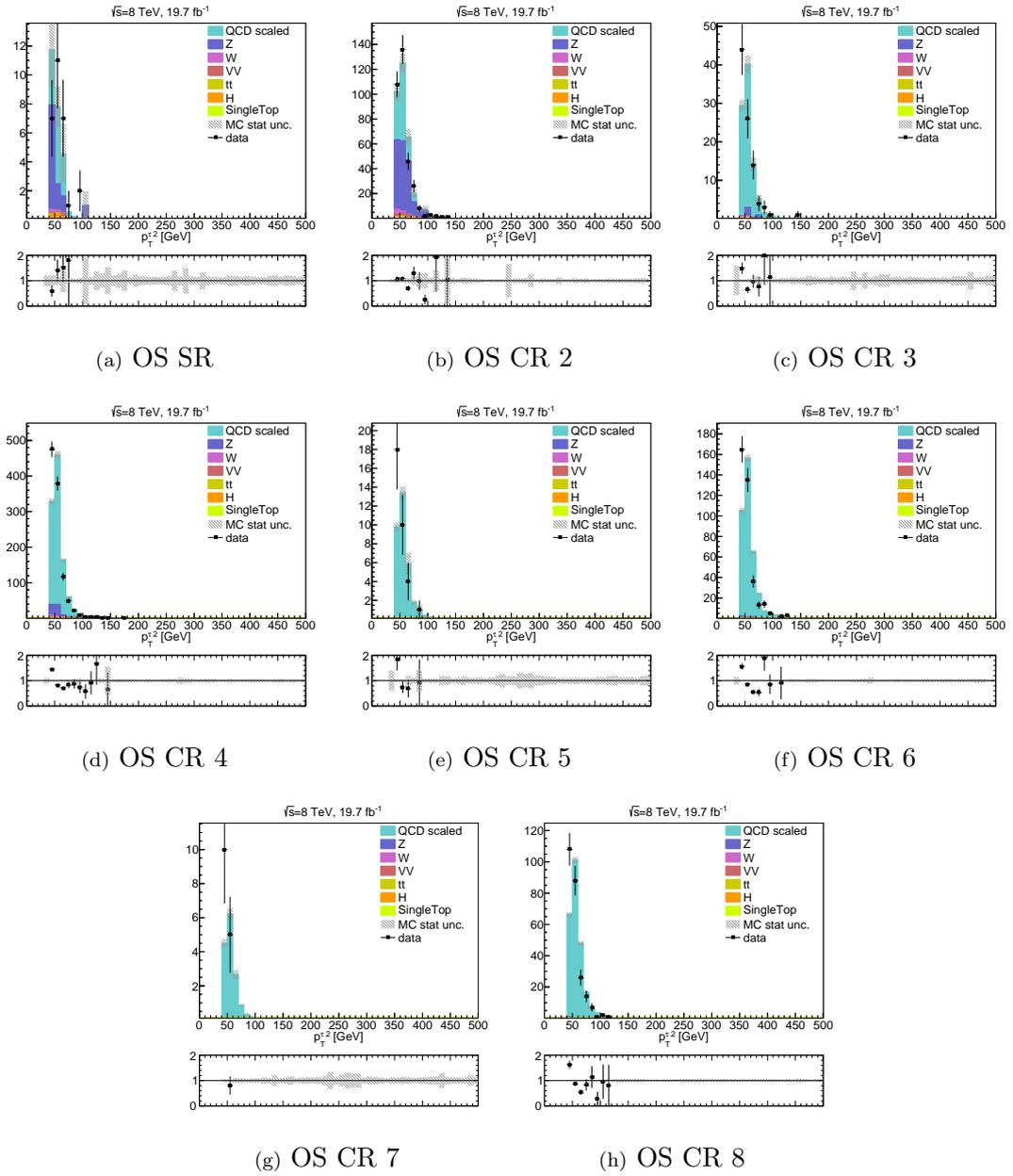
FIGURE B.81: Linear second to leading τ transverse momentum for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
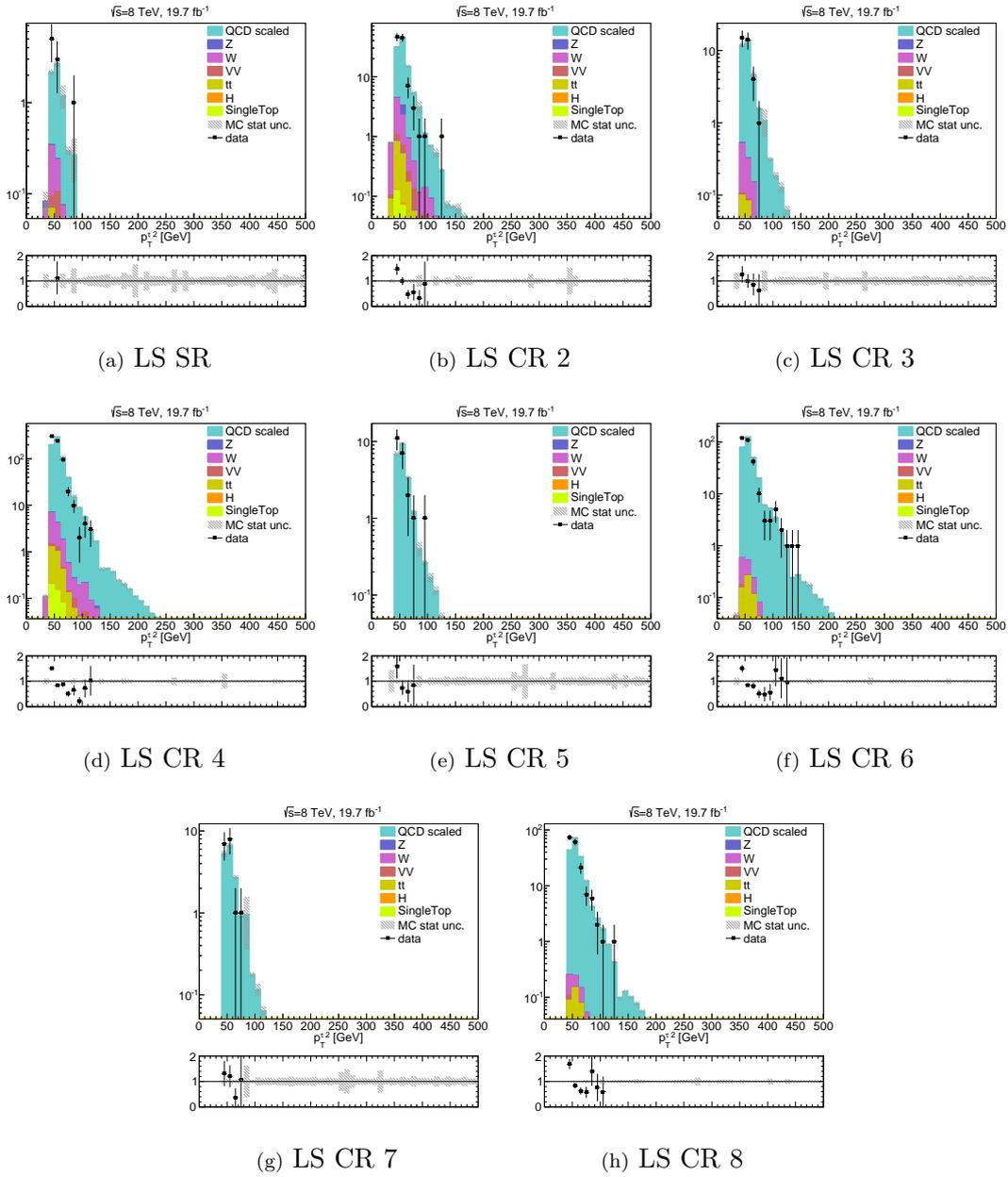
FIGURE B.82: Logarithmic second to leading τ transverse momentum for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
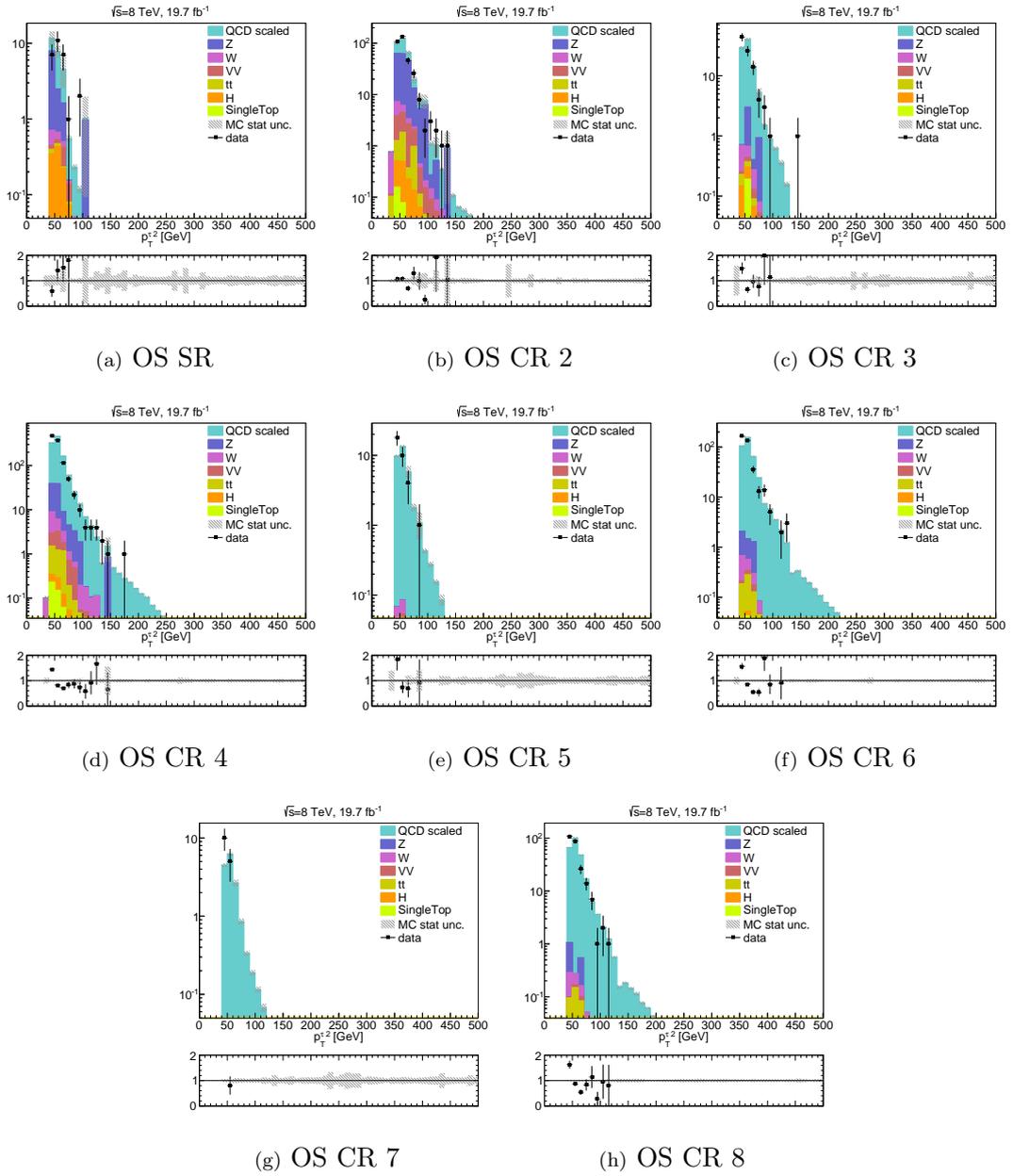
FIGURE B.83: Logarithmic second to leading τ transverse momentum for OS

B.4.21 $\eta^{\tau\ 1}$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
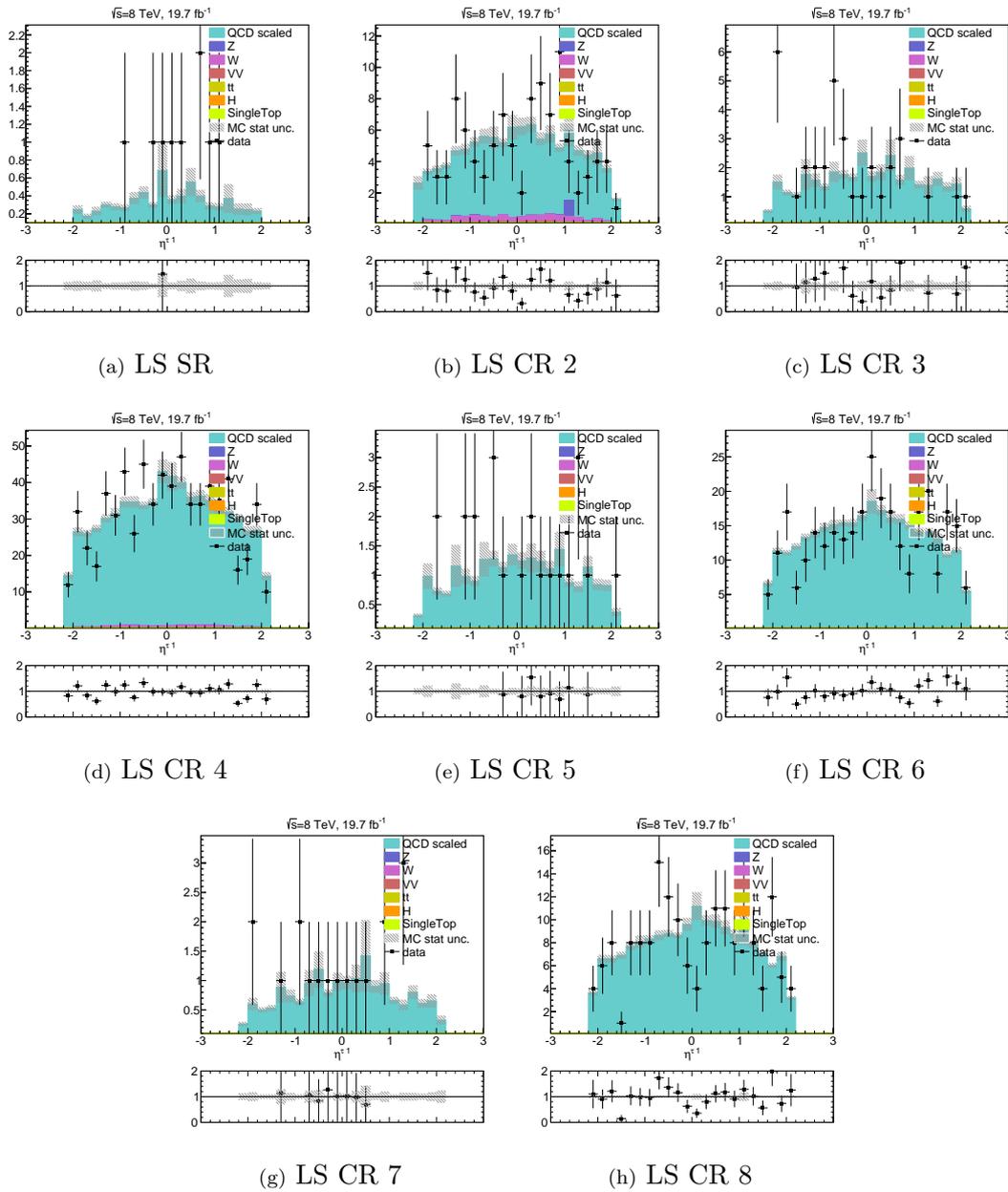
FIGURE B.84: Linear leading τ pseudorapidity for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
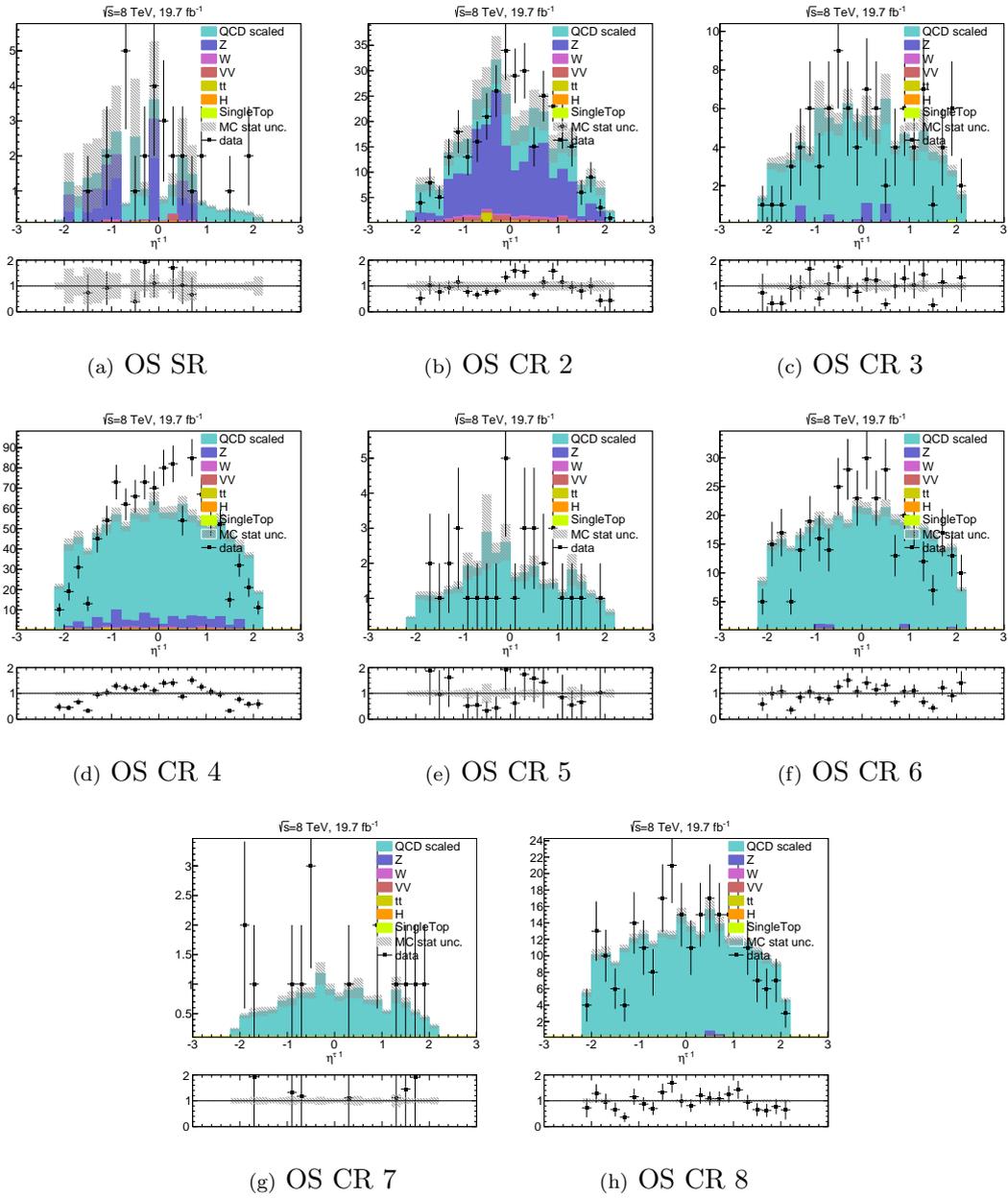
FIGURE B.85: Linear leading τ pseudorapidity for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
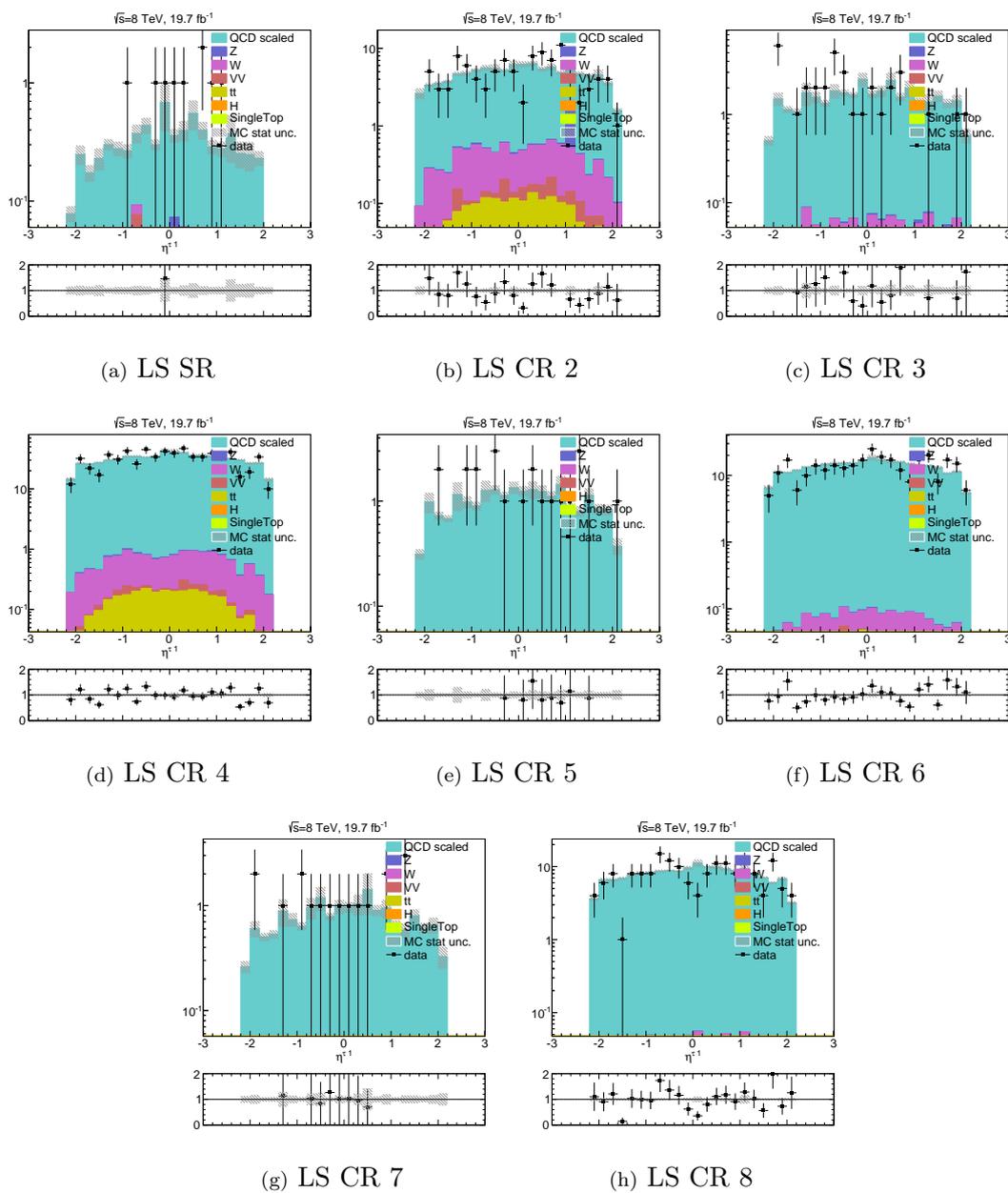
FIGURE B.86: Logarithmic leading τ pseudorapidity for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
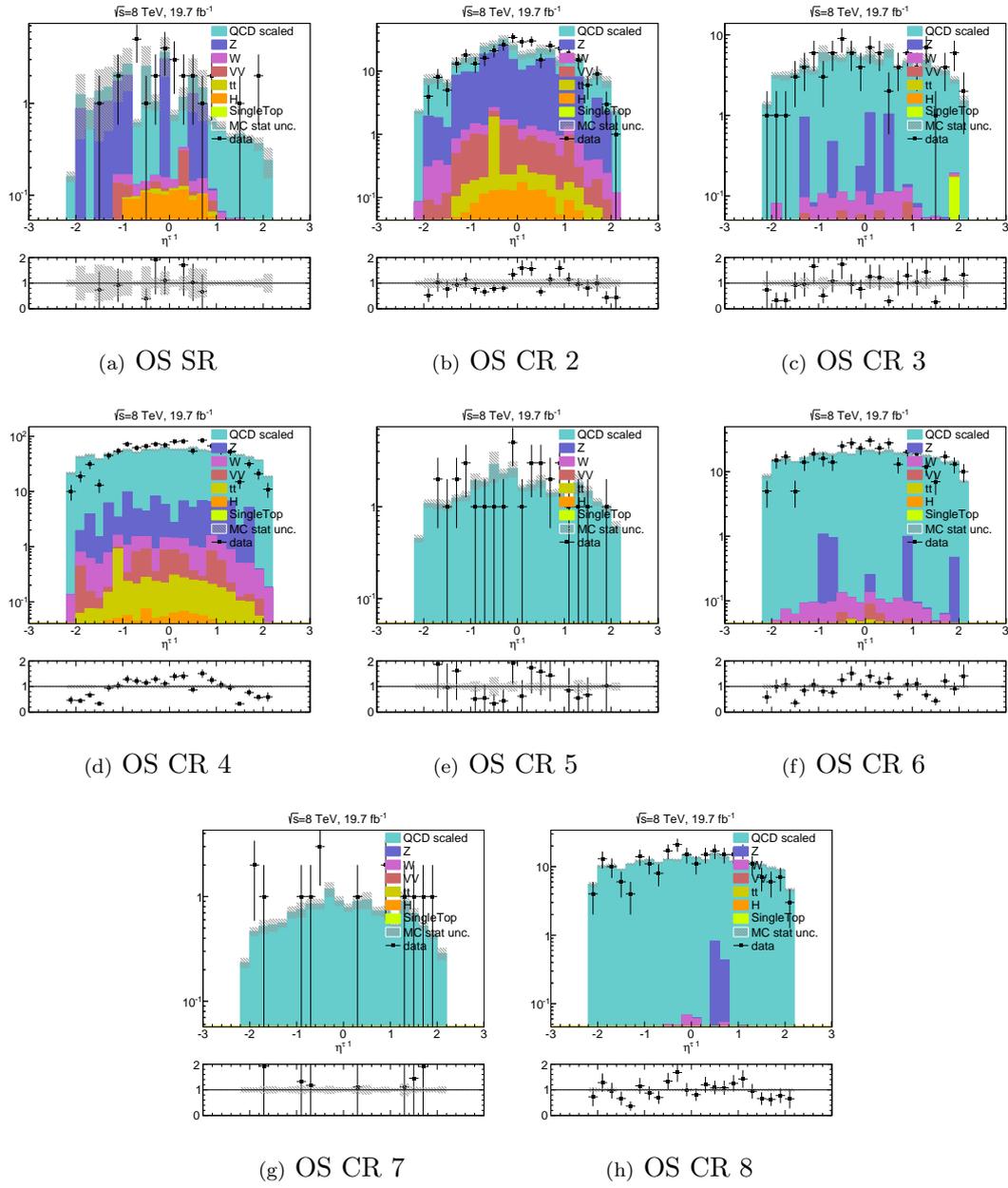
FIGURE B.87: Logarithmic leading τ pseudorapidity for OS

B.4.22 $\eta^{\tau\,2}$

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
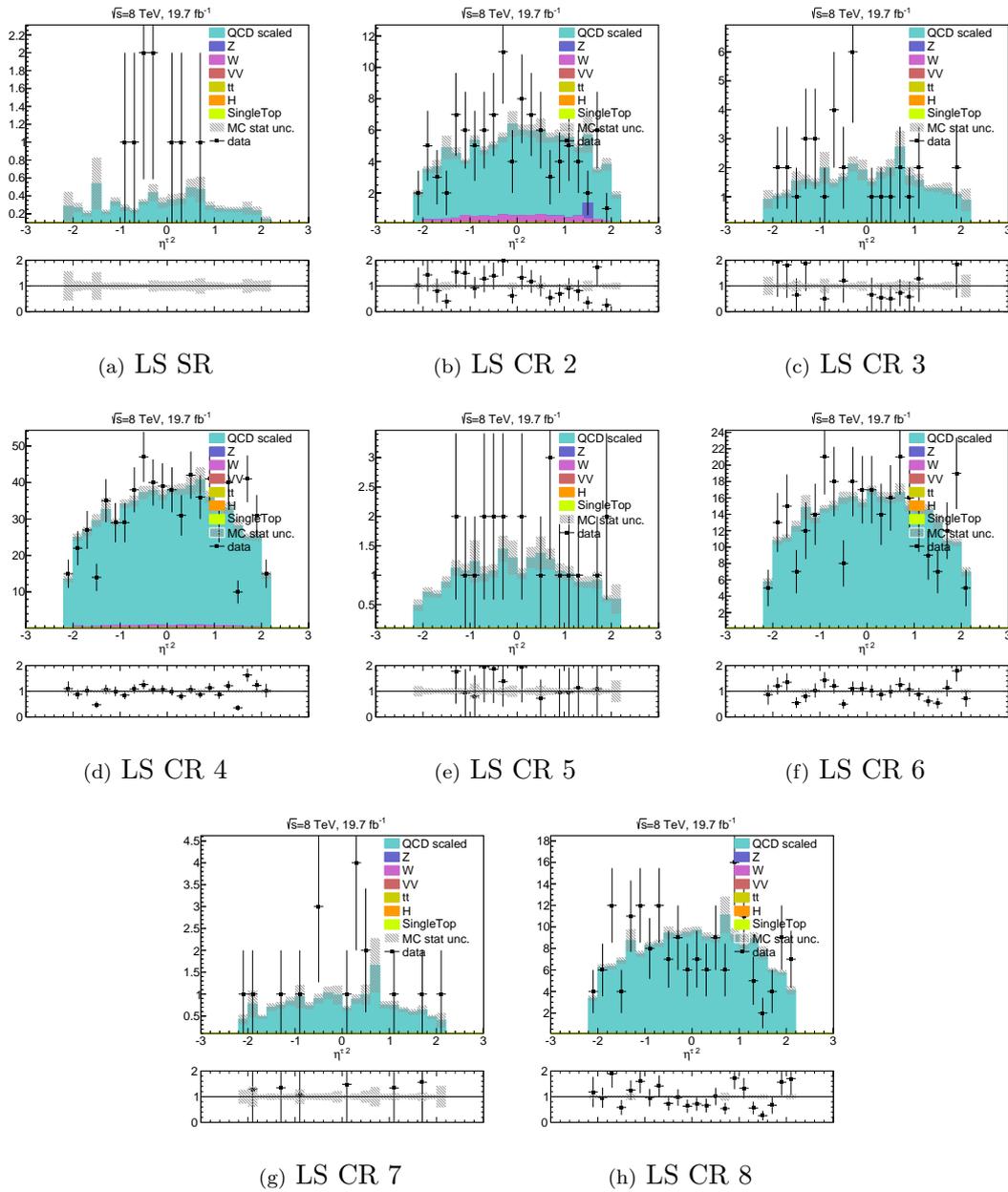
FIGURE B.88: Linear second to leading τ pseudorapidity for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
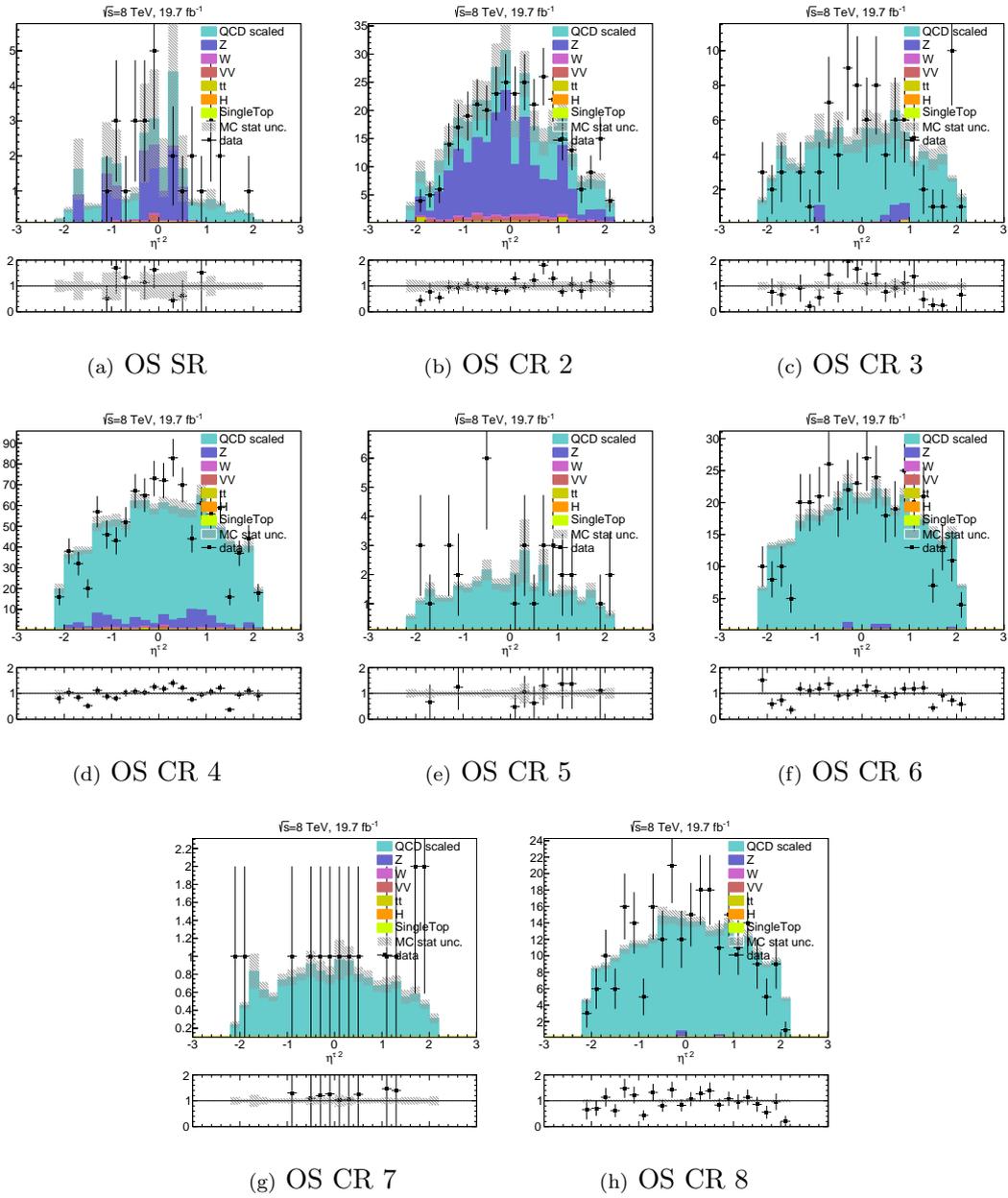
FIGURE B.89: Linear second to leading τ pseudorapidity for OS

(a) LS SR

(b) LS CR 2

(c) LS CR 3

(d) LS CR 4

(e) LS CR 5

(f) LS CR 6

(g) LS CR 7

(h) LS CR 8
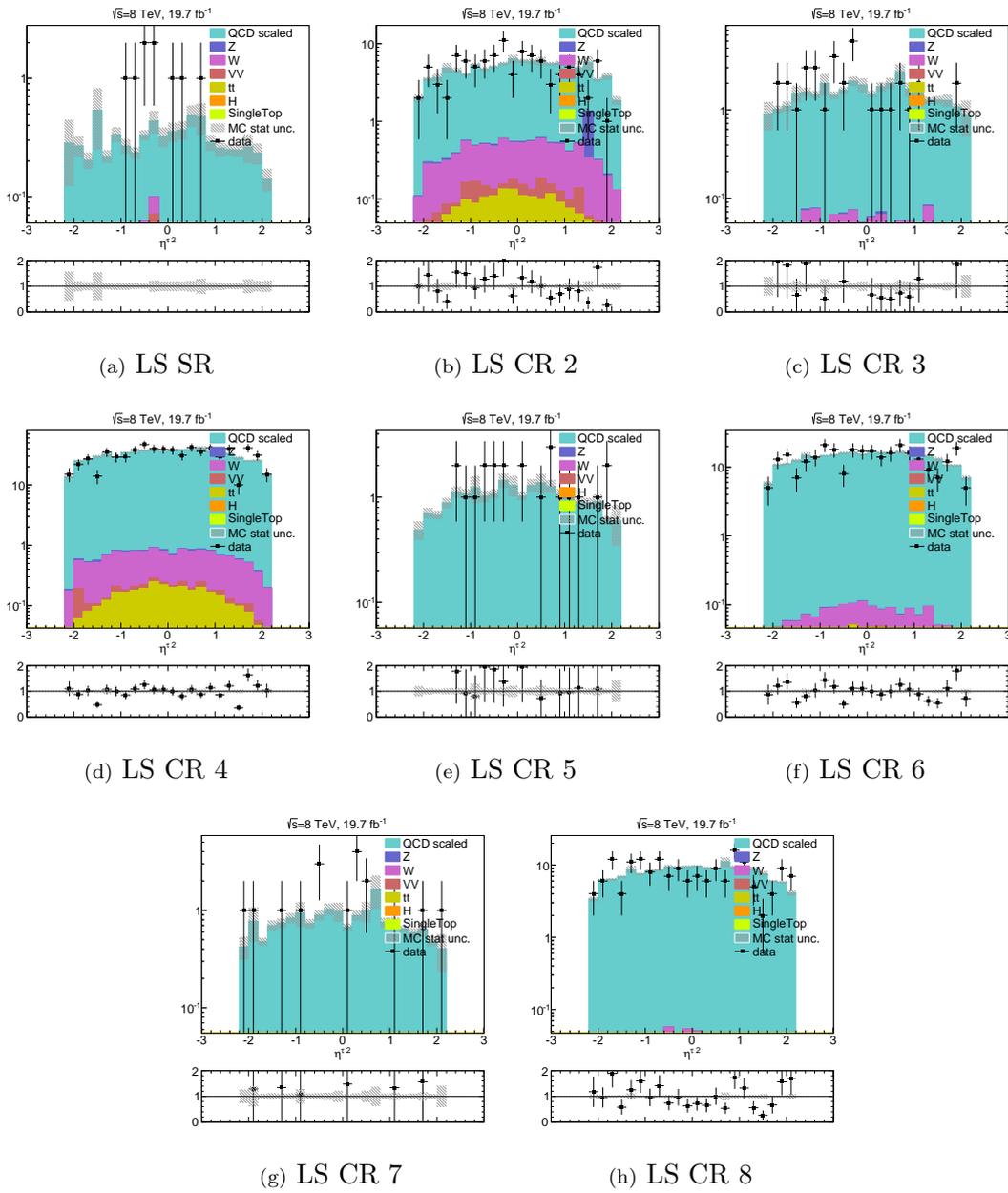
FIGURE B.90: Logarithmic second to leading τ pseudorapidity for LS

(a) OS SR

(b) OS CR 2

(c) OS CR 3

(d) OS CR 4

(e) OS CR 5

(f) OS CR 6

(g) OS CR 7

(h) OS CR 8
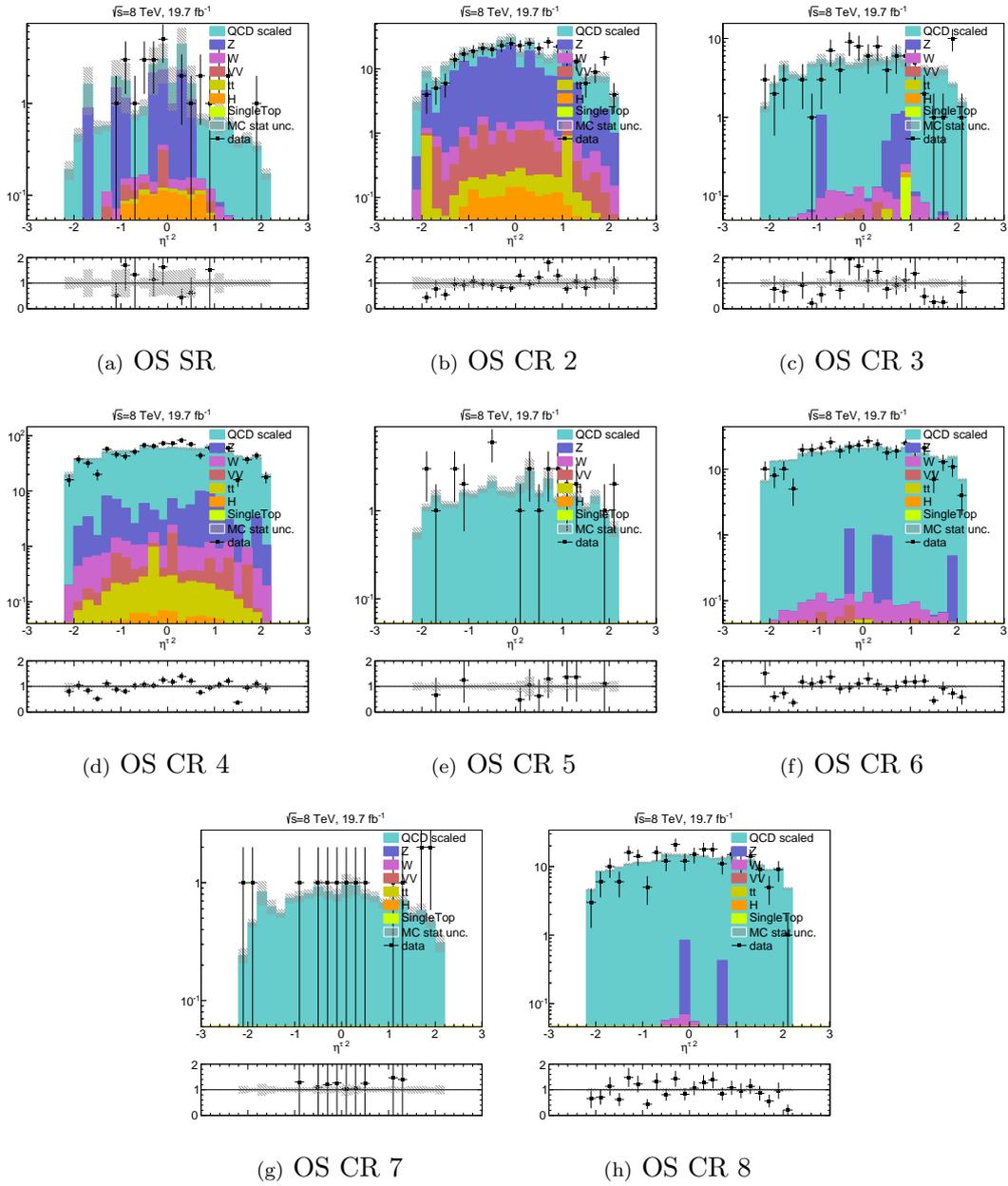
FIGURE B.91: Logarithmic second to leading τ pseudorapidity for OS

Abbreviations

ALICE	**A L**arge **I**on **C**ollider **E**xperiment
APD	**A**valanche **P**hoto**D**iode
ATLAS	**A T**oroidal **LHC A**pparatu**S**
BNV	**B**aryon **N**umber **V**iolation
BSM	**B**eyond the **S**tandard **M**odel of particle physics
CERN	**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire
CMS	**C**ompact **M**uon **S**olenoid
CL	**C**onfidence **L**evel
CR	**C**ontrol **R**egion
CSC	**C**athode **S**trip **C**hamber
EB	**E**CAL **B**arrel
EC	**E**nd**C**aps
EE	**E**CAL **E**ndcaps
ECAL	**E**lectromagnetic **CAL**orimeter
EM	**E**lectro**M**agnetic
FSR	**F**inal **S**tate **R**adiation
HB	**H**adronic **B**arrel
HBHE	**H**adronic **B**arrel **H**adronic **E**ndcaps
HCAL	**H**adronic **CAL**orimeter
HE	**H**adronic **E**ndcaps
HF	**H**adronic **F**orward
HLT	**H**igh-**L**evel **T**rigger
HPD	**H**ybrid **P**hoto**D**iode
HPS	**H**adron **P**lus **S**trip algorithm
ISR	**I**nitial **S**tate **R**adiation

IT	**I**n **T**ime
JEC	**J**et **E**nergy **C**orrections
JER	**J**et **E**nergy **R**esolution
JES	**J**et **E**nergy **S**cale
LHC	**L**arge **H**adron **C**ollider
LNV	**L**epton **N**umber **V**iolation
L1	**L**evel **1** (trigger)
LS	**L**ike **S**ign $\rightarrow$ sign$(q_{\tau_1} \cdot q_{\tau_2}) = +1$
MET	**M**issing **E**nergy in the **T**ransverse plane
MSSM	**M**inimal **S**uper**S**ymmetric **S**tandard **M**odel
OOT	**O**ut **O**f **T**ime
OS	**O**pposite **S**ign $\rightarrow$ sign$(q_{\tau_1} \cdot q_{\tau_2}) = -1$
PDF	**P**arton **D**istribution **F**unction
PF	**P**article **F**low
PFchs	**P**article **F**low with **c**harged **h**adron **s**ubstraction
PS	**P**arton **S**hower
PU	**P**ile**U**p
QCD	**Q**uantum **C**hromo**D**ynamics
$\overline{\textbf{QCD}}$	inverted **Q**uantum **C**hromo**D**ynamics event selection
QED	**Q**uantum **E**lectro**D**ynamics
QFD	**Q**uantum **F**lavour**D**ynamics
RBX	**R**eadout **B**o**X**
RD	**R**un **D**ependent
RMS	**R**oot **M**ean **S**quare
SM	**S**tandard **M**odel of particle physics
SR	**S**ignal **R**egion
SUSY	**SU**per**SY**mmetry
TEC	**T**racker **E**nd**C**aps
TIB	**T**racker **I**nner **B**arrel
TID	**T**racker **I**nner **D**isks
TOB	**T**racker **O**uter **B**arrel
UE	**U**nderlying **E**vent
VBF	**V**ector **B**oson **F**usion

$\overline{\textbf{VBF}}$ inverted **V**ector **B**oson **F**usion event selection

VEV   **V**acuum **E**xpectation **V**alue

List of Figures

List of Tables

Bibliography

[1] Search for electroweak SUSY with VBF tagging. Technical Report CMS-PAS-SUS-14-005, CERN, Geneva, 2015.

[2] Ian Stewart Jack Cohen, Terry Pratchett. *The Science of Discworld IV: Judgement Day*. Random House, 2013.

[3] CERN hackfest. Standard model infographic. http://www.isgtw.org/spotlight/go-particle-quest-first-cern-hackfest, august 2012.

[4] David Griffiths. *Introduction to Elementary Particle Physics*. Wiley-VCH, Portland OR, USA, 2010.

[5] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys.Rev.*, 145:1156–1163, 1966. doi: 10.1103/PhysRev.145.1156.

[6] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[7] Xerxes Tata Howard Baer. *Weak Scale Supersymmetry*. Cambridge University Press, Cambridge CB2 2RU, UK, 2006.

[8] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, Joey Huston, Jon Pumplin, et al. Implications of CTEQ global analysis for collider observables. *Phys.Rev.*, D78:013004, 2008. doi: 10.1103/PhysRevD.78.013004.

[9] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur.Phys.J.*, C63:189–285, 2009. doi: 10.1140/epjc/s10052-009-1072-5.

[10] J.W.F. Valle. Status and implications of neutrino masses: a brief panorama. 2015.

[11] Yupeng Yang. Constraints on the basic parameters of dark matter using the Planck data. 2015.

[12] G. Scelza and A. Stabile. Numerical Analysis of Galactic Rotation Curves. 2015.

[13] A.L. King, T.M. Davis, K. Denney, M. Vestergaard, and D. Watson. High Redshift Standard Candles: Predicted Cosmological Constraints. *Mon.Not.Roy.Astron.Soc.*, 441:3454–3476, 2014. doi: 10.1093/mnras/stu793.

[14] K. Abe et al. Search for proton decay via $p \to \nu?K^+$ using 260 kilotonyear data of Super-Kamiokande. *Phys.Rev.*, D90(7):072005, 2014. doi: 10.1103/PhysRevD.90.072005.

[15] CMS collaboration. Summary of cms susy results in sms framework. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS, ICHEP 2014.

[16] LHC SUSY Cross Section Working Group. Susy theory cross section plots. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSYSMSSummaryPlots8TeV, april 2015.

[17] Michael Kramer, Anna Kulesza, Robin van der Leeuw, Michelangelo Mangano, Sanjay Padhi, et al. Supersymmetry production cross sections in *pp* collisions at $\sqrt{s} = 7$ TeV. 2012.

[18] Marcela Carena, Stefania Gori, Nausheen R. Shah, Carlos E.M. Wagner, and Lian-Tao Wang. Light Stau Phenomenology and the Higgs $\gamma\gamma$ Rate. *JHEP*, 1207:175, 2012. doi: 10.1007/JHEP07(2012)175.

[19] Search for direct EWK production of SUSY particles in multilepton modes with 8TeV data. Technical Report CMS-PAS-SUS-12-022, CERN, Geneva, 2013. URL `https://cds.cern.ch/record/1546777`.

[20] Bhaskar Dutta, Alfredo Gurrola, Will Johns, Teruki Kamon, Paul Sheldon, et al. Vector Boson Fusion Processes as a Probe of Supersymmetric Electroweak Sectors at the LHC. *Phys.Rev.*, D87(3):035029, 2013. doi: 10.1103/PhysRevD.87.035029.

[21] Thomas Sven Pettersson and P Lefevre. The Large Hadron Collider: conceptual design. Technical Report CERN-AC-95-05-LHC, CERN, Geneva, Oct 1995.

[22] Mike Lamont. Status of the lhc. *Journal of Physics: Conference Series*, 455(1):012001, 2013. URL `http://stacks.iop.org/1742-6596/455/i=1/a=012001`.

[23] S. Beole et al. ALICE technical design report: Detector for high momentum PID. 1998.

[24] A. Airapetian et al. ATLAS: Detector and physics performance technical design report. Volume 1. 1999.

[25] A. Airapetian et al. ATLAS: Detector and physics performance technical design report. Volume 2. 1999.

[26] G.L. Bayatian et al. CMS physics: Technical design report. 2006.

[27] G.L. Bayatian et al. CMS technical design report, volume II: Physics performance. *J.Phys.*, G34:995–1579, 2007. doi: 10.1088/0954-3899/34/6/S01.

[28] 629958. LHCb technical design report: Reoptimized detector design and performance. 2003.

[29] Christiane Lefevre. The CERN accelerator complex. Complexe des accelerateurs du CERN. Dec 2008.

[30] Eike Steffen Schlieckau. Measurement of Top-Quark Mass and Inclusive Top-Quark-Pair Production Cross Section in pp Collisions at $\sqrt{s} = 7/8$ TeV with CMS.

[31] CMS Collaboration. CMS Luminosity Public Results, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. March 2013.

[32] John R Hubbard. LHC trigger design. 1997.

[33] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[34] Martin Gorner. Differential Cross Sections for Top-Quark-Pair Production in the e/μ+Jets Final State at $\sqrt{s} = 8$ TeV in CMS.

[35] D. Sprenger, M. Weber, R. Adolphi, R. Brauer, L. Feld, et al. Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap. *JINST*, 5:P06007, 2010. doi: 10.1088/1748-0221/5/06/P06007.

[36] 2012 ECAL detector performance plots. Mar 2013. URL `https://cds.cern.ch/record/1528235`.

[37] G. Baiatian et al. Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges. 2007.

[38] Victor Daniel Elvira. Measurement of the Pion Energy Response and Resolution in the CMS HCAL Test Beam 2002 Experiment. Technical Report CMS-NOTE-2004-020, CERN, Geneva, Sep 2004.

[39] Johan Alwall, Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet, et al. MadGraph/MadEvent v4: The New Web Generation. *JHEP*, 0709:028, 2007. doi: 10.1088/1126-6708/2007/09/028.

[40] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 1006:043, 2010. doi: 10.1007/JHEP06(2010)043.

[41] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. doi: 10.1088/1126-6708/2006/05/026.

[42] Serguei Chatrchyan et al. Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV. *JHEP*, 1109:109, 2011. doi: 10.1007/JHEP09(2011)109.

[43] M. Bahr, S. Gieseke, M.A. Gigg, D. Grellscheid, K. Hamilton, et al. Herwig++ Physics and Manual. *Eur.Phys.J.*, C58:639–707, 2008. doi: 10.1140/epjc/s10052-008-0798-9.

[44] Stefan Gieseke, Christian Rohr, and Andrzej Siodmok. Colour reconnections in Herwig++. *Eur.Phys.J.*, C72:2225, 2012. doi: 10.1140/epjc/s10052-012-2225-5.

[45] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506: 250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[46] 1230052. Comparison of the Fast Simulation of CMS with the first LHC data. 2010.

[47] Ahmed Ali and Gustav Kramer. Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD. *Eur.Phys.J.*, H36: 245–326, 2011. doi: 10.1140/epjh/e2011-10047-1.

[48] Denis Rathjens. Particle Reconstruction at the LHC using Jet Substructure Algorithms.

[49] Gavin P. Salam. Towards Jetography. *Eur.Phys.J.*, C67:637–686, 2010. doi: 10.1140/epjc/s10052-010-1314-6.

[50] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur.Phys.J.*, C72:1896, 2012. doi: 10.1140/epjc/s10052-012-1896-2.

[51] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008. doi: 10.1088/1126-6708/2008/04/063.

[52] Francesca Cavallari. Performance of calorimeters at the LHC. *J.Phys.Conf.Ser.*, 293:012001, 2011. doi: 10.1088/1742-6596/293/1/012001.

[53] Erika Garutti. Calorimeters - energy measurement. http://www.desy.de/ garutti/LECTURES/ParticleDetectorSS12/L10_Calorimetry.pdf, february 2015.

[54] Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. Technical Report CMS-PAS-PFT-10-002, CERN, Geneva, 2010.

[55] E. Tziaferi N. Saoulidou. Performance of the particle-flow jet identification criteria using proton-proton collisions at $\sqrt{s} = 8$ tev. *CMS analysis note*, 2014. URL `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2014_227_v4.pdf`.

[56] b-Jet Identification in the CMS Experiment. Technical Report CMS-PAS-BTV-11-004, CERN, Geneva, 2012.

[57] Electron reconstruction and identification at sqrt(s) = 7 TeV. Technical Report CMS-PAS-EGM-10-004, CERN, Geneva, 2010.

[58] Single Muon efficiencies in 2012 Data. Mar 2013.

[59] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012. doi: 10.1088/1748-0221/7/10/P10002.

[60] Study of tau reconstruction algorithms using pp collisions data collected at sqrt(s) = 7 TeV. Technical Report CMS-PAS-PFT-10-004, 2010.

[61] Serguei Chatrchyan et al. Performance of tau-lepton reconstruction and identification in CMS. *JINST*, 7:P01001, 2012. doi: 10.1088/1748-0221/7/01/P01001.

[62] MET performance in 8 TeV data. Technical Report CMS-PAS-JME-12-002, CERN, Geneva, 2013.

[63] The CMS collaboration. Jet energy scale and resolution in the 8 tev pp data. *CMS paper (to be published)*, 1(8):8–41, December 2014. URL `http://cms.cern.ch/iCMS/analysisadmin/versions?analysis=JME-13-004`.

[64] Serguei Chatrchyan et al. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011. doi: 10.1088/1748-0221/6/11/P11002.

[65] Determination of the Relative Jet Energy Scale at CMS from Dijet Balance. Technical Report CMS-PAS-JME-08-003, CERN, 2009. Geneva, May 2009.

[66] University of Hamburg CMS group. Kalibri framework. *CMS TWiki https://twiki.cern.ch/twiki/bin/view/CMS/HamburgWikiAnalysisCalibration and http://www.iexp.desy.de/groups/cms/kalibri/api/index.html*, 2007-2015.

[67] *Determination of the jet energy scale in CMS*, volume 404, 2012. doi: 10.1088/ 1742-6596/404/1/012013.

[68] Jet Energy Corrections and Uncertainties. Detector Performance Plots for 2012. Jul 2012.

[69] Status of the 8 TeV Jet Energy Corrections and Uncertainties based on 11 fb^{-1} of data in CMS. May 2013.

[70] *Jet Energy Calibration in the CMS experiment*, volume 60, 2013. doi: 10.1051/ epjconf/20136020010.

[71] 8 TeV Jet Energy Corrections and Uncertainties based on 19.8 fb^{-1} of data in CMS. Oct 2013.

[72] Matthias Schroder. Quality of jet measurements and impact on a search for new physics at CMS. pages 88–90.

[73] Kristin Goebel. Probing supersymmetry based on precise jet measurements at the CMS experiment.

[74] A. C. Florez Bustos A. Gurrola A.K. Kalsi T. Kamon D. Marconi S. Maruyama K. Mazumdar D. Rathjens C. Sander J. Singh V. Bhatangar, N. Dhingra and L. Vanelderen. Search for vbf produced ew gauginos in events with 2 hadronically decaying τ-leptons. *CMS analysis note*, 1(8):1–109, August 2014. URL `http://cms. cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2014/099`.

[75] O. Bouhali A. Celik J. Cumalat C. Davis A. Delannoy N. Dhingra W. Flanagan A. Florez A. Gurrola W. Johns A. Kalsi T. Kamon D. Kim E. Luiggi S. Malik D. Marconi S. Maruyama K. Mazumdar A. Melo R. Montalvo M. Narain Y. Oh D. Rathjens C. Sander P. Sheldon J. Singh C. Avila, V. Bhatangar and L. Vanelderen. Search for ewkinos produced in vector boson fusion processes with $\sqrt{s} = 8$ tev. *CMS analysis note*, 1(8):1–56, August 2012. URL `http://cms.cern.ch/iCMS/jsp/db_ notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2012/321`.

[76] W Adam, V Adler, B Hegner, L Lista, S Lowette, P Maksimovic, G Petrucciani, S Rappoccio, F Ronga, R Tenchini, and R Wolf. PAT: the CMS Physics Analysis Toolkit. Technical Report CMS-CR-2009-083, CERN, Geneva, May 2009.

[77] CMS Collaboration. https://twiki.cern.ch/twiki/bin/viewauth/CMS/UserCodeNWUntupleProducer. January 2014.

[78] Victor Mukhamedovich Abazov et al. Jet energy scale determination in the D0 experiment. *Nucl.Instrum.Meth.*, A763:442–475, 2014. doi: 10.1016/j.nima.2014. 05.044.

[79] Wolfgang Ochs and Redamy Perez Ramos. Particle Multiplicity in Jets and Sub-jets with Jet Axis from Color Current. *Phys.Rev.*, D78:034010, 2008. doi: 10.1103/PhysRevD.78.034010.

[80] Sebastian Naumann-Emme. Simultaneous measurement of top quark mass and jet energy scale using template fits at the CMS experiment.

List of Publications

The following publications resulted from the work performed in this thesis:

- D. Rathjens *et al.* [CMS Collaboration], "Jet Energy Calibration in the CMS experiment", EPJ Web Conf. **60** (2013) 20010

- To be published: The CMS collaboration, "Jet Energy Scale and Resolution in 8 TeV pp data", JME-13-004

- The CMS collaboration, "Search for electroweak SUSY with VBF tagging", CMS-PAS-SUS-14-005 [1], to be submitted to JHEP

Acknowledgements