SUMMER STUDENT PROJECT:
N-Jettiness as a Jet Multiplicity Measurement

D. Damgaard
Supervised by:
M. Seidel & P. C. Harris

September 1, 2017

Introduction

The N-Jettiness [1] event shape, developed by I. Stewart, F. Tackmann, and W. J. Waalewijn and included in the Fastjet library is a method to veto for how N-jet like an event looks, returning a value to each number of jets N.
N-Jettiness is used in NNLO calculations see for instance [2]. Minimizing the jettiness for each N, we might get an estimate of number of jets in an event.

The main method of calculating the observable in question consists of minimizing the geometrical distance between charged particle 4 - momenta and axises defined by various methods. In our case the axises are determined by XCone. The jettiness observable is returned as the summed transverse momenta, P_t, which is left over after the jets are found.

In this summer student project, developed July and August 2017, $t\bar{t}$-events are generated using Pythia [3] within the Rivet [4] framework.

Theory

In particle scattering, at high enough energy, there is potential to generate QCD objects like quark or gluons. Since these SU(3) objects are "confined" at lower energy, they are not discovered directly, but generate jets of hadrons, like protons, pions, or even kaons. The "jets" or particle showers are observables in most detectors like CMS, ATLAS, or ALICE.

We are interested in an "event shape", an observable we can use to understand the geometrical properties of a certain event such as how many jets we have in said event. A candidate for this method proposed by I. Stewart et al. is N-Jettiness. A variable calculated by the following:

$$\tau_N = \sum_k \min\left(\rho(q_a, p_k), \rho(q_b, p_k), \rho(q_1, p_k), \cdots, \rho(q_N, p_k)\right)$$

Where k runs over every final particle in the event (that are not signal leptons or photons) and $\rho(q,p)$ is the distance between the jet particles and the proposed "jet-axis", q, which are defined by various methods. In this project the geometrical measure used is the XConeMeasure defined in [5]. The N-Jettiness observable, τ_N retained from the XConeMeasure can be interpreted as the summed P_t not associated with N jets and thus tells us how well the event is being described by having N jets.

Since this value is trivially minimized for $N \to \#$ of charge final state particles, we expect that a simple minimization method will not be enough to retain an estimate of the event shape, thus two methods are proposed:

1. Cut method: this method is a minimization scheme which minimizes the τ_N(N) down to a certain cut - This produces a large overflow since the ranges of τ_N(N) is very different from event to event
2. Slope method: This method is a maximization of $|\frac{d\tau_N}{d\mathrm{N}}|$

The main event we are considering is the decay of a pair of top quarks ($t\bar{t}$) as seen on Fig(2) which mainly decay into b-quarks, an electron, a muon + missing transverse energy due to neutrinos. This is due to the large element in the CKM-matrix between top and bottom-quarks. Due to the short width of the tau, it decays instantaneously to its products in comparison to the resolution of modern detectors, we are not able to detect them directly. The b-quarks however have a larger width compared to the resolutions and thus travels a short distance before decaying. This means that the rest frame of the decaying b-quarks is not at the interaction vertex and thus a discriminant for b-quarks are found [6]. The b-quarks themselves decay to their lighter cousins, who in turn hadronize and deposit their energy in the detector calorimeter.

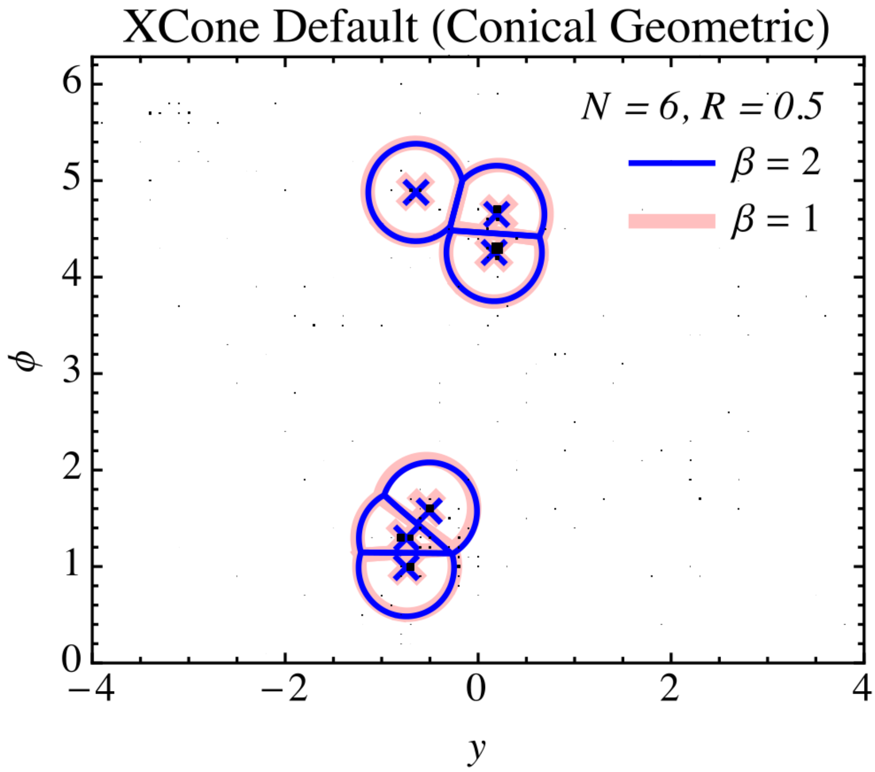

Figure 1: XCones for N = 6 in $t\bar{t}$- fully hadronic events, τ_6 would be the summed P_t outside the rings

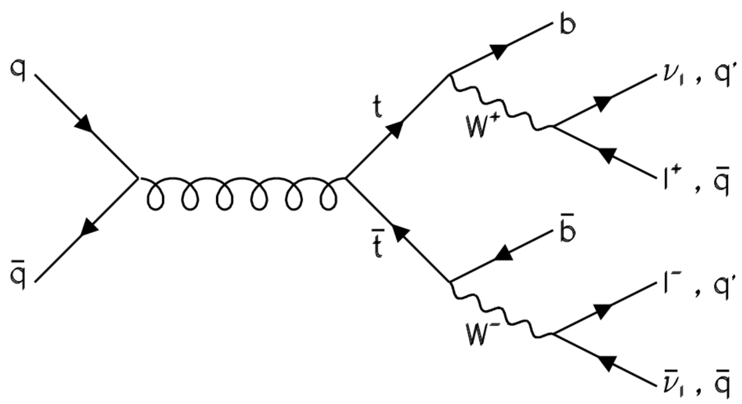

Figure 2: The $t\bar{t}$- decay. The channel of interest has the W^+/W^- decays to an electron and a muon, this channel is very clean and provides good selection in data

Studying the observables using RIVET

First of foremost the events are generated. Most of this is done through RIVET in the CMS working environment, while the parton shower is done in PYTHIA.
Simulations were done with ISR scale variation in the RIVET framework in order to predict higher order effects, this is shown in Fig(4), and τ_2, τ_4, τ_6, and τ_8 are plotted together on Fig(3) in order to show how the $\tau_N \to 0$ when $N \to \infty$ as described in section (2)

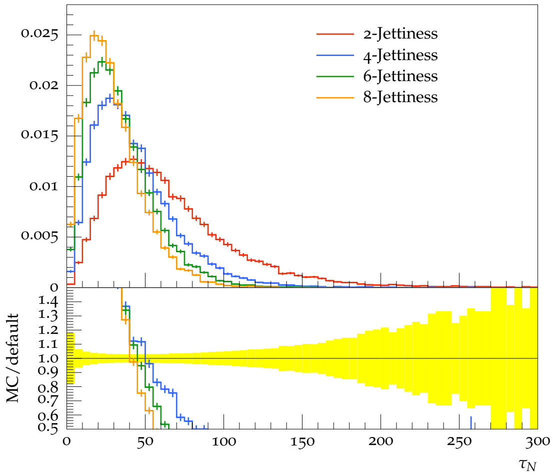

Figure 3: The 2-,4-,6-, and 8-Jettiness plotted against each other to show how higher N distribution tends toward 0.

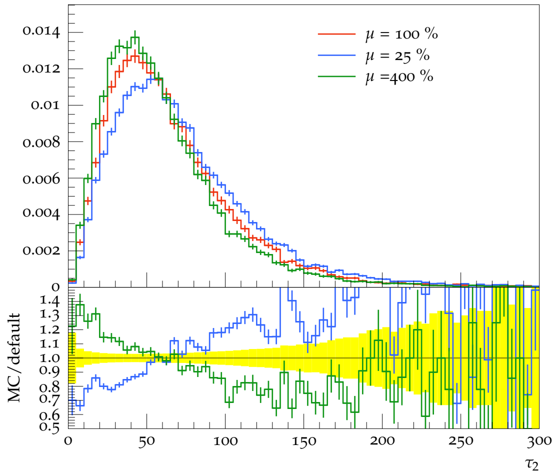

Figure 4: τ_2-Jettiness on detector level for $\mu = (0.25, 1.00, 4.00)$ we note higher scale forces τ_2 towards lower values as α_{QCD} rises

The Jettiness value is calculated and stored, this is later referred to as generator- or particle level N-jettiness. Parton showers are simulated using PYTHIA. We also simulate detector effects and store our events as reconstructed- or detector level.

Variations of ISR and FSR scales are also simulated for systematics, as well as a sample using the HERWIG parton shower instead of PYTHIA.

Analysis

Event selection

Since we are studying $t\bar{t}$-decay in the CMS detector [7], we choose the following cuts:

- Lepton transverse momenta: $P_t \geq 25$ GeV
- Invariant mass of lepton product: $M \geq 12$ GeV
- Lepton pseudo-rapidity: $|\eta| \leq 2.5$
- Two b-tagged jets with transverse momentum $P_t \geq 30$ GeV
- Leptons $\supset e, \mu$

Furthermore only charged final state hadrons are considered in order to keep a the channel clean.

Signal/Background after event selection

After applying the cuts, the SM background is negligible and the data is almost completely described by MC signal. Thus we neglect background samples in further analysis.

N-Jettiness Observable on detector level

As described in the theory section, the N-Jettiness observable have the tendency to lower as N rises. This is in accordance with what we expect as τ_N measures the summed P_t that is not associated with N jets and is trivially minimized when N is equal to the number of charged particles in final state.

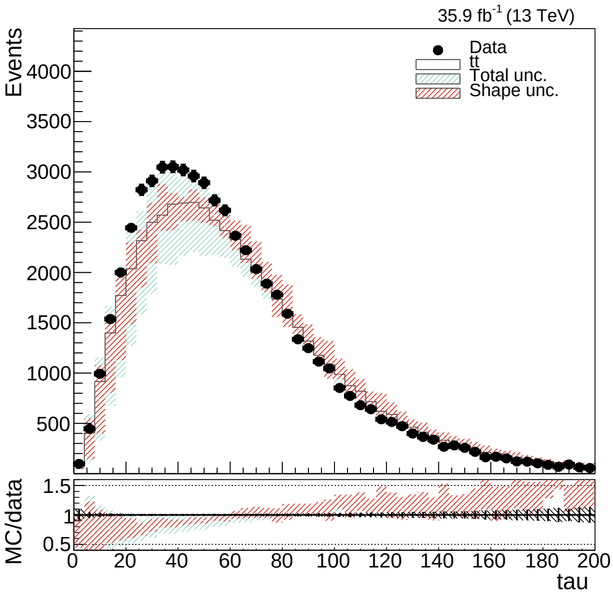

Figure 5: τ_2-Jettiness at detector level

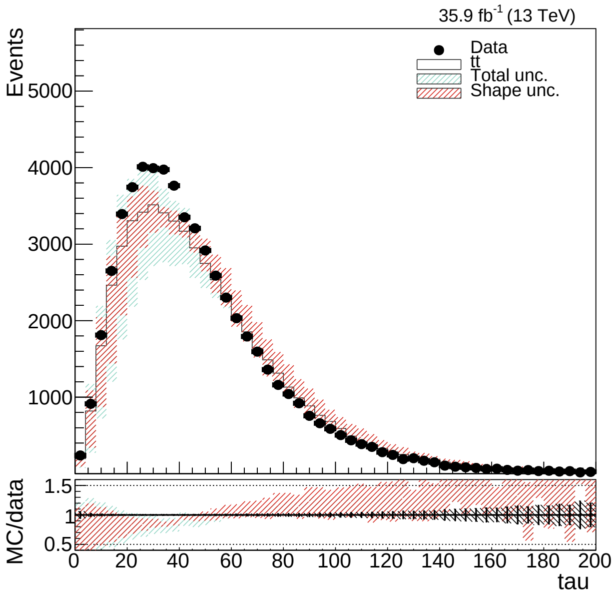

Figure 6: τ_3-Jettiness at detector level

On figures Fig(5) and Fig(6) we see the data plotted together with the $t\bar{t}$-simulated events with systematic and statistical errors. The data is not described completely by the systematic, since we are not taking all possible sources into account e.i. tracking uncertainty or jet energy scales.
The Slope and Cut method histograms are shown on figures Fig(7) and Fig(8). The overflow of the Cut method is placed in bin 9.

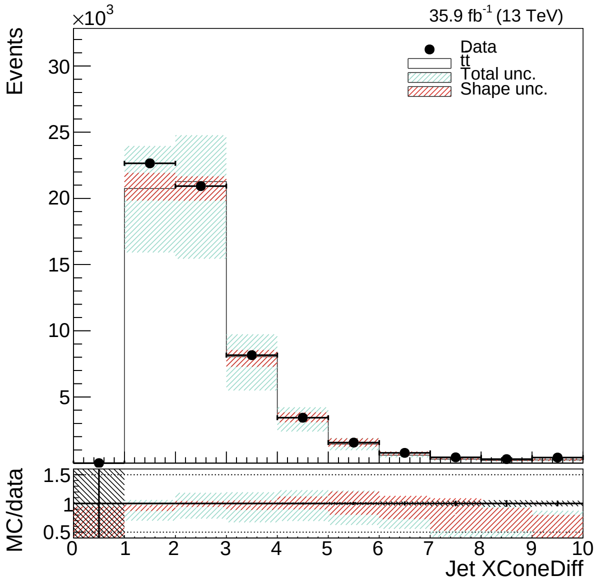

Figure 7: Slope method at detector level

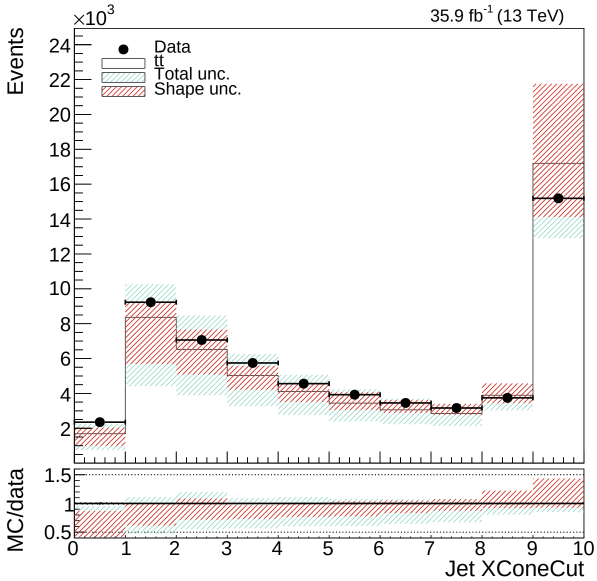

Figure 8: Cut method at detector level

Migration and Unfolding

For the MC samples we have the N-jettiness observable at both detector and particle level. These are plotted in a 2d histograms and the binning is optimized. The optimization is done by normalizing columns, which gives the purity in the diagonal and corresponds to the migration in to the given bin when reconstructing to detector level, and normalizing the rows, which gives the stability in diagonal and corresponds to the migration out of the bin. We combine bins until we have in each bin purity ≥ 50 % and stability ≥ 40 %. Before and after optimization of τ_2 is shown on figures Fig(9) and Fig(10).

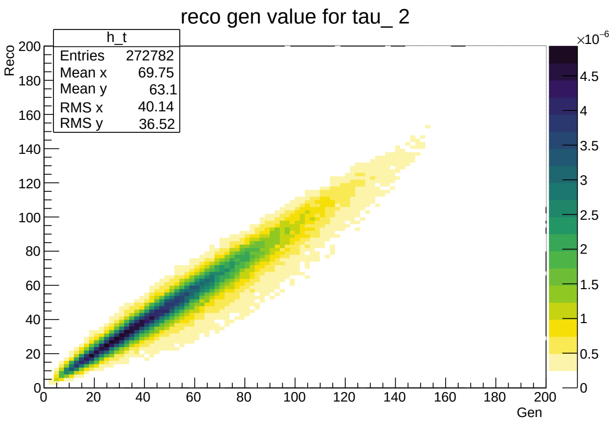

Figure 9: Before Optimization

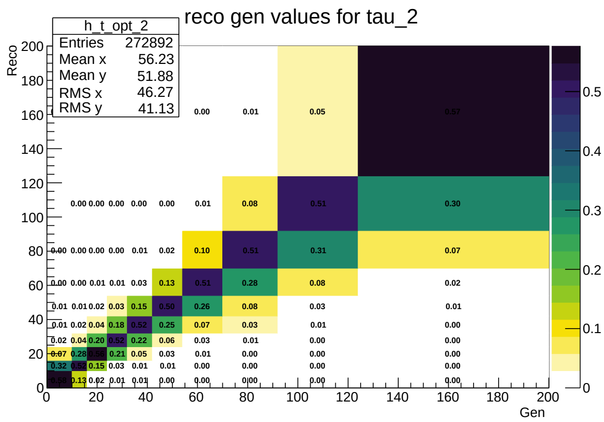

Figure 10: After Optimization

This was done for the following observables: τ_N - Jettiness, for $N = 0, \cdots 9$ (Fig(13), Fig(14),(A)), jet multiplicity using the cut Fig(8) and slope Fig(7), as well as the conventional NJ algorithm Fig(15).

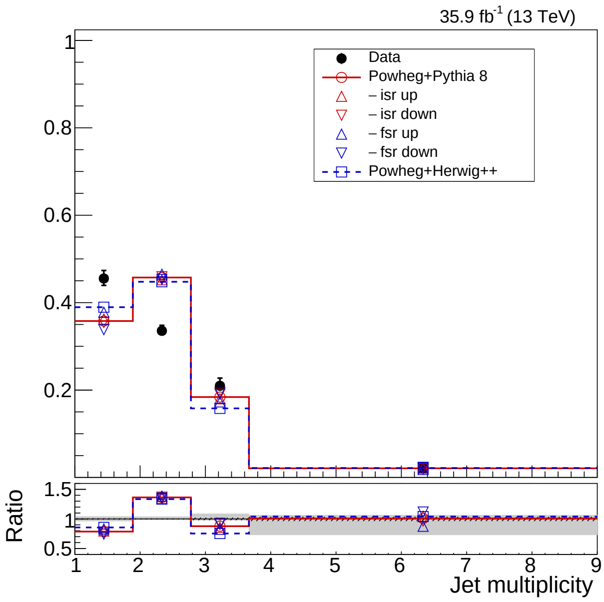

Figure 11: Slope jet multiplicity unfolded (The regularization might not refrain)

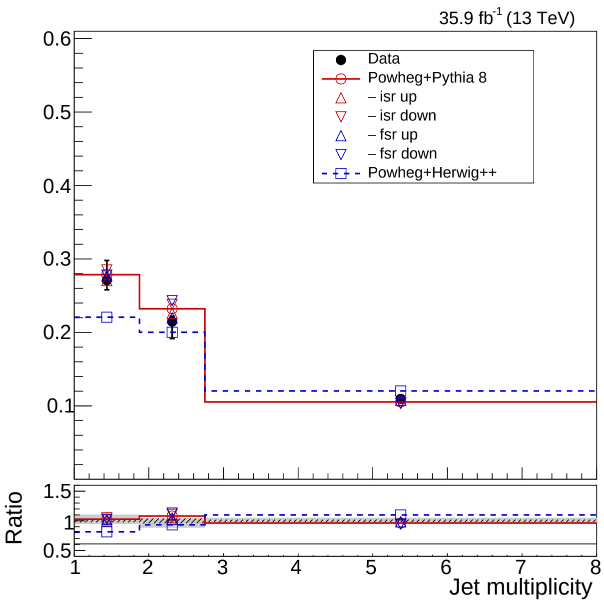

Figure 12: Cut jet multiplicity unfolded

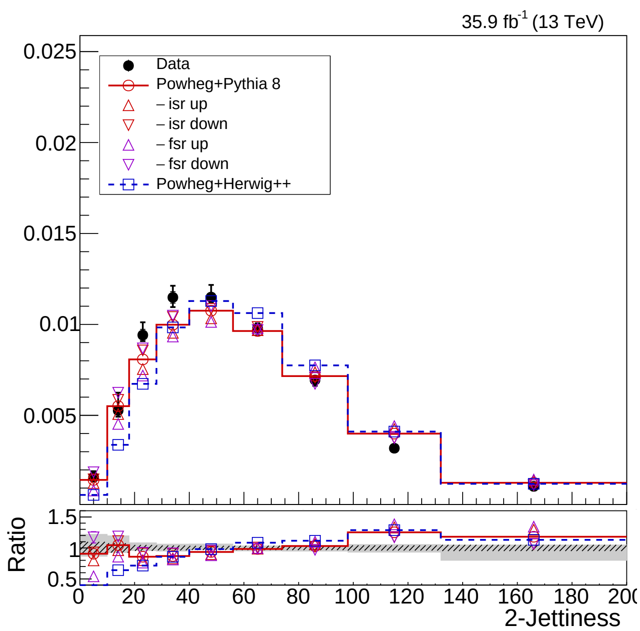

Figure 13: τ_2- Jettiness Unfolded

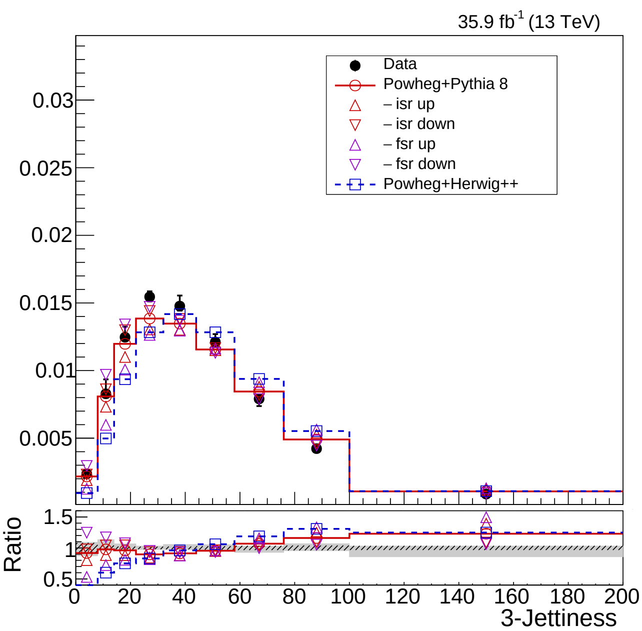

Figure 14: τ_3- Jettiness Unfolded

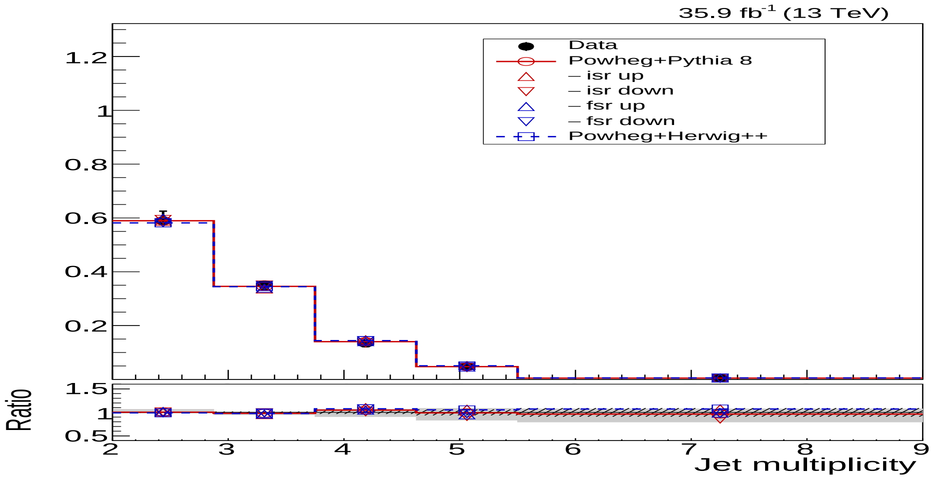

Figure 15: NJ Jet Multiplicity

Unfolding of the data was performed on every jettiness and multiplicity histogram by the ROOT library TUnfold using an optimized regularization parameter. Systematics are added to the Monte Carlo samples.
We note how neither of observables are particularly well described by Monte Carlo aside from NJ even with the included systematics. This might be a sign of New Physics or not understood effects or, they might disappear when more systematics are included. A very pronounced example of this can be seen on Fig(11) where we see the first bin is less populated that the second bin in the MC, which is expected according to our event selection, however the data is a continuously falling distribution, interestingly the Herwig and Pythia parton showering disagree wildly with in themselves and the data, this is in stark contrast with the NJ Fig(15) where the data and both parton showers seems to be in agreement.

Conclusion

N-jettiness have been studied at detector level. The Monte Carlo samples have been used to produce migration matricies, which in turn, have been used to unfold the data to particle level, and thus be compared to theory.
The unfolded data is shown not to be well described by the MC generator with the two different parton showers.

Two methods for estimating the jet multiplicity is introduced and studied both at detector and particle level together with the conventional jet algorithm NJ. Here we see a large discrepancy between data and samples and even between the showering models themselves for the proposed methods, while NJ is very well described and the models seem to be in agreement. This tells us that as of yet, we do not have a one-to-one replacement of NJ and more work is needed to be conducted on understanding the parton showering models and the N-jettiness distributions.

Appendix A

We plot N-Jettiness from N = $0 \cdots 9$:

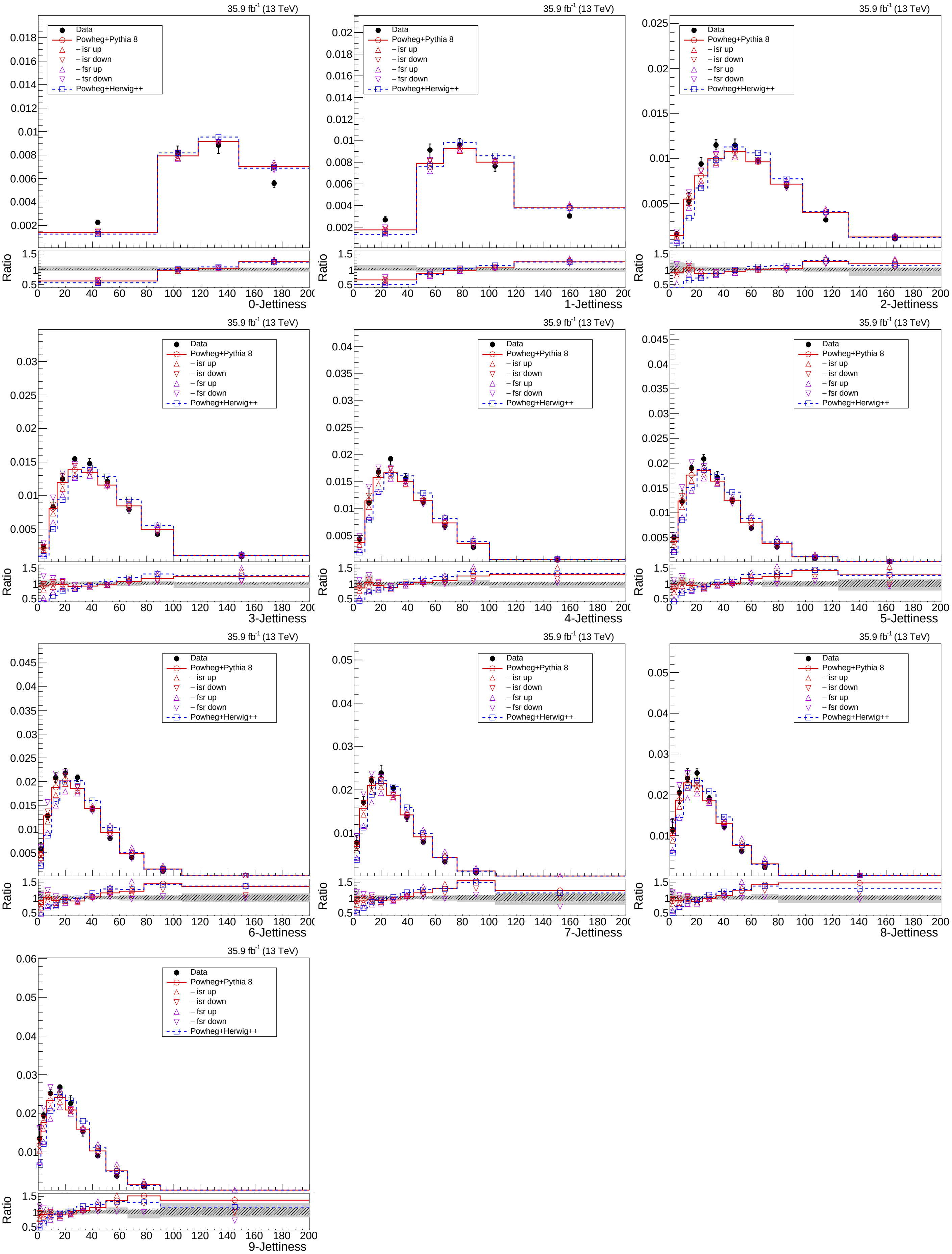

Figure 16: N-jettiness from $N \cdots 0 - 9$

References

[1] Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. N-Jettiness: An Inclusive Event Shape to Veto Jets. *Phys. Rev. Lett.*, 105:092002, 2010.

[2] Jonathan Gaunt, Maximilian Stahlhofen, Frank J. Tackmann, and Jonathan R. Walsh. N-jettiness Subtractions for NNLO QCD Calculations. *JHEP*, 09:058, 2015.

[3] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[4] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. Rivet user manual. *Comput. Phys. Commun.*, 184:2803–2819, 2013.

[5] Iain W. Stewart, Frank J. Tackmann, Jesse Thaler, Christopher K. Vermilion, and Thomas F. Wilkason. XCone: N-jettiness as an Exclusive Cone Jet Algorithm. *JHEP*, 11:072, 2015.

[6] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.

[7] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

,