Tuning of the Whizard Monte Carlo event generator via the Professor tool

Philipp Englert, TU Dortmund, Germany

September 10, 2015

Abstract

This work explains the essentials of the systematic tuning procedure that is used by the Professor tool. We attempt to reproduce the Pythia6 tune, that was done 2009 by the Professor development team to get a better understanding of the effects of different ways of executing the tuning procedure on the results and to be able to tune the Whizard shower method confidently. Within the limits of the uncertainties of the tune, this attempt was successful, but because of the large uncertainties of some of the tune parameters and the fact that some tune results are located outside of the ranges of our parameter sampling, the results are not fully trustworthy and we advise to repeat the tunes with adapted parameter ranges and a larger amount of parameter points. The analysis of the tune of the Whizard shower showed even larger errors and therefore, we advise to revise this tune too.

Contents

1 Introduction **3**

2 Theory **3**

2.1 Parametrisation . . . 3

2.2 Goodness of fit function . . . 4

2.3 Factorisation of parameter sets . . . 5

3 Tuning of the Whizard event generator **6**

4 Comparisons and summary **12**

4.1 Different longitudinal fragmentation functions . . . 12

4.2 QED initial state radiation . . . 13

4.3 Photon emissions included in the shower evolution, no QED initial state radiation . . . 13

4.4 The effect of the number of parameter points on the results . . . 14

4.5 Whizard tune . . . 16

5 Appendix **18**

5.1 Comparison plots . . . 18

5.1.1 PYTHIA shower . . . 18

5.1.2 WHIZARD shower . . . 24

1 Introduction

Both in theoretical and experimental particle physics, Monte Carlo event generators are of particular importance. They can be used to analyse experiments or to make predictions, that are useful to design future experiments. The necessity of Monte Carlo simulators arises from the complexity of the final states observed in experiments and the non-perturbative aspects of QCD at low energies [1].
The main goal of Monte Carlo simulators is to describe the reality correctly and every physical model that accomplishes this, contains free parameters like, coupling constants. Furthermore, Monte Carlo event generators have to make use of approximations of the considered physical models. Therefore, the best fitted values of the exact values of the contained parameters do not necessarily yield the best description of observables, when compared to experiments. We are therefore looking for a set of effective values of these parameters, such that the event generator in question describes the reality best. The process of finding this set is called tuning and can be systematized, which is e.g. done by the tuning tool PROFESSOR [2]. The subject of my project in the DESY theory group this summer was to look into the process of tuning the event generator WHIZARD [3] and as part of that, to tune for the first time the analytic parton shower together with PYTHIA hadronisation in a systematic way.

2 Theory

The PROFESSOR tuning system provides a way to systematically tune the parameters of Monte Carlo event generators to experimental data as opposed to the traditional, labour-intensive guesswork of manual tuning. The idea is as follows: We sample a sufficient amount of points in parameter space and run our generator once for each of these points. The results of the event generations are used to calculate a selection of observables, we are interested in. Afterwards, we parametrise these observables bin-by-bin, dependent on our physical parameters and calculate a goodness of fit function (GoF function), yet to be defined. Finally, we maximise this GoF measure and check, if our result is actually a global maximum of this measure.
Of course, the above is a sloppy description of the tuning process and the following treatise will describe the difficulties, that have been neglected so far, in more detail.

2.1 Parametrisation

Our first task is to sample points in parameter space. To do so, we need to specify a range of possible parameters and a hypercube $[\boldsymbol{p}_{\min}, \boldsymbol{p}_{\max}]$ seems convenient. The choice of boundaries requires physics input, since they should include all thinkable values but they should also be as narrow as possible so that the parametrisations of the observables describe the real generator response accurately in the hypercube without needing too many sample points [2].
The minimum number of samples needed depends on the type of parametrisation of the

observables. An n-th order polynomial in P-dimensional parameter space requires at least

$$N_n^{(P)} = 1 + \sum_{i=1}^{n} \frac{1}{i\,!} \prod_{j=0}^{i-1} (P + j) \tag{1}$$

sample points [2]. In the special case of a second order polynomial, which will be our first choice, the required number is:

$$N_2^{(P)} = 1 + P + P(P+1)/2\,. \tag{2}$$

The reason, why we choose to parametrise the observable bins, rather than the GoF function, is the following. We expect the generator response to be much more complicated than the second order polynomial, which is why the parametrisation will be inaccurate, but the errors of the descriptions of the individual observable bins should be overall decorrelated. Therefore, we hope, that these errors cancel out to some extent, when we calculate the GoF function [2].
As another way to account for the complexity of the real generator response, we do not only sample the minimum amount N_2^P of needed parameter points, but a significantly bigger amount $N_{\text{tune}} \gg N_2^P$. This enables us to use a more sophisticated method of determining the coefficients of the polynomials, which represents a least squares fit to the true generator response. This method is called the pseudoinverse and is implemented numerically in the Professor tool. As a rule of thumb, N_{tune} should be at least twice as big as N_2^P [2, 4].

2.2 Goodness of fit function

One of the most difficult tasks of tuning an event generator is the choice of observables to tune to. This mostly depends on the purpose of the event generator. It can't possibly describe every thinkable observable well, so one has to decide which ones are the most important.
A big advantage of interpolating every observable bin is, that one can assign weights w_{b_O} to them individually, when taken into account in the GoF function. The Professor system uses a χ^2 function as a GoF measure and minimises it. It is defined as

$$\chi^2(\boldsymbol{p}) = \sum_{O} \sum_{b_O} w_{b_O} \frac{(f^{(b_O)}(\boldsymbol{p}) - R_{b_O})^2}{\Delta_{b_O}^2}\,, \tag{3}$$

where O indicates the observable and b_O the corresponding bins, $f^{(b_O)}$ is the fit function of the bin b_O and R_{b_O} the corresponding reference data and Δ_{b_O}. The total uncertainty of the experimental reference is Δ_{b_O} and is taken into account, such that bins with a big uncertainty are of small significance to the fit. In the numerical implementation of the goodness of fit function, bins with $\Delta_{b_O} = 0$ are excluded, because it usually indicates, that there is no reference data at all for the bin in question [2].

A rule of thumb for choosing the observable weights w_O is to assign a smaller weight to observables that belong to a bigger set of observables with similar physical content. Furthermore, bigger weights can be assigned to observables that are not described well enough by the tune, which is relevant for reruns of the generator. It is also necessary to pay attention, that the minimisation result is not too sensitive to the choice of weights. But overall, there is no such thing as a strict rule to choose the observable weights except the purpose of the tuning itself, which is to describe reality as good as possible.
Finally, the GoF function has to be maximised or, as in the case of PROFESSOR , the χ^2 function has to be minimised. It is important to make sure that the result of the minimisation is not only a local minimum of the χ^2 function but also the global minimum, which is not a trivial task because of the complexity of the GoF function.
Another aspect, we have to bear in mind, is that we might have chosen not enough parameter points for the parametrisation. One way to make up for this is to oversample even a bit more, perform tunes for several distinct subsets N_{tune} of the total sample N and compare them amongst each other. One should make sure that the subsets are sufficiently decorrelated, which is why one should usually oversample by approximately 30% [2].

2.3 Factorisation of parameter sets

Given a general purpose event generator, we want to tune a large amount P of parameters which might not seem feasible on first sight because the volume of the parameter space grows exponentially with P and the amount of sample points should grow similarly so that the parametrisations of the observables represent a good description of the true generator response for the whole parameter space [2]. A good approach to solve this problem is to factorise the set of tune parameters $M = \{p_i\}$ in decoupled subsets $N_i \subset M$. In this case, decoupled means that the minimisation result of the GoF measure $\chi^2(N_1, N_2, N_3, \ldots)$ with respect to the parameter set N_1 has to be independent of the values of the parameters that belong to N_2, N_3, ... and vice versa. The most obvious case of such a decoupling would be that the observables that are sensitive to the parameters N_1 are insensitive to N_2, N_3, ..., etc. [5].
In fact, it is not necessary that the parameters decouple in exactly the sense, described above, but it is sufficient, if e.g. only the minimisation with respect to N_1 is independent of the values of all of the parameters N_2, N_3, ... as opposed to the minimisation of N_2, which may depend on the values of N_1 and be only independent of N_3, ... and so forth. In that specific case, the tuning can be parted into several successive stages: In the first tuning stage, N_1 is tuned, afterwards N_2, whilst using the tuned values of the parameters N_1 and so on. Further factorization of the subsets N_i is an option, but only if the requirements are satisfied.
As a consequence, the tuning of an event generator is normally parted into three successive stages. First, the final state parameters, i.e. final-state shower and hadronization parameters are tuned to e^+e^- observables, since they should be universal between lepton and hadron collisions and event generators do not need to use initial-state parton shower parameters to describe the processes in question. Once tuned, the final state parameters

can be used in the tuning process of the initial-state shower tuning, which is normally the second stage. Afterwards, multi-parton-interaction parameters and beam remnant effects are tuned [5].

3 Tuning of the Whizard event generator

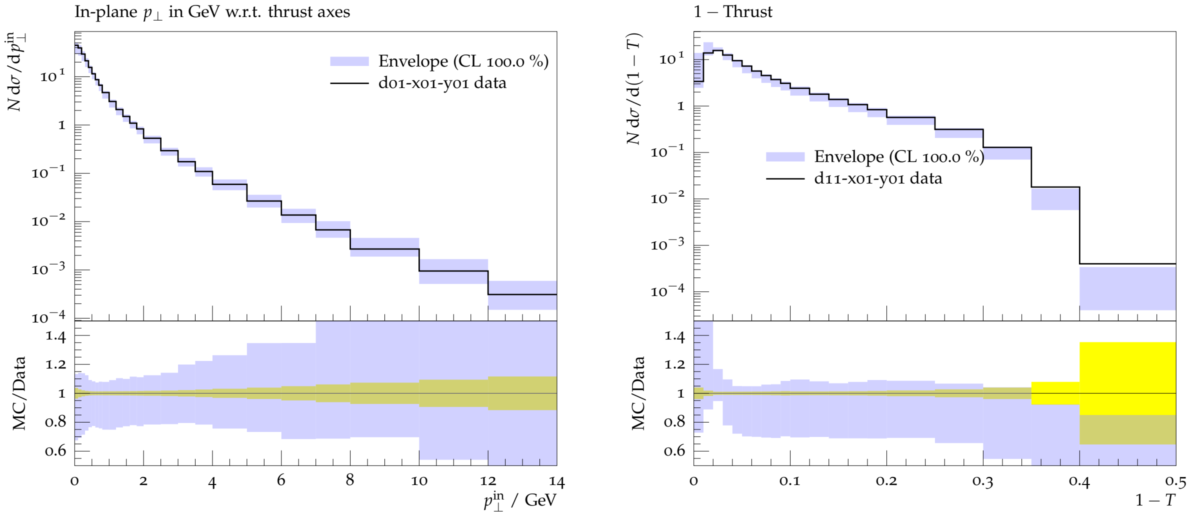
The focus of this section is the first stage of the tuning of the WHIZARD event generator with the PROFESSOR system. The parameters, which are tuned, are the coupling constant of the strong final state interactions $\alpha_{\mathrm{s,f}}$, the width of the transverse momentum distribution σ_q, the a and b parameter of the Lund symmetric fragmentation function, the Bowler modification factor for b-Quarks r_b and the shower cut-off. The WHIZARD system provides different ways of handling parton showers: One is to use the internal implementation of WHIZARD , the other one is to use an interface that delegates the job to PYTHIA . The tune has to be performed for both shower methods. The hard process, that is used for the tunes, is

$$e^+ + e^- \to q + \bar{q}\,.$$

The analysis of the generated events is done with the RIVET tool [6], which calculates observables of the desired analysis based on the generator output. We consider an analysis that is already included in RIVET , namely the `DELPHI_1996_S3430090` [7]. It features event shape and charged particle inclusive distributions, which were determined by analysing measurements of decays of the Z-boson into hadrons, as it was measured by the DELPHI detector at LEP.

Furthermore, for each shower method, three tunes have to be performed, each of them representing a different choice of longitudinal fragmentation function: the Lund symmetric fragmentation function (1), the Lund symmetric fragmentation function for heavy endpoint quarks, modified according to the Bowler space-time picture of string evolution (2) and a variation of the latter, which makes it possible to interpolate between Bowler and Lund separately for c and b quarks (3). The main goal of this work is to understand the systematic tuning process and the issues that will inevitably come up. Therefore, the first step is to try to reproduce the results of an established tune of the PYTHIA shower method. The tune (3) was performed with and without consideration of photon emission in the shower evolution and in the former case with and without consideration of QED initial state radiation. In the following, I will show the tuning process for the case (3) of the above-mentioned with the PYTHIA shower method and without any considerations of photon emission exemplarily and afterwards summarise and compare the results of the other tunes.

The first step of the tuning procedure is to produce envelope plots, which show the range of the observables that can be calculated, based on the generated events, compared to the reference data of the considered analysis. These plots represent an easy way to check, if the scan ranges in parameter space are reasonable for the upcoming interpolation. Two examples of the produced envelope plots are shown in Figure 1. The former shows $1/N \mathrm{d}\sigma/\mathrm{d}p_\perp^{\mathrm{in}}$ plotted against $p_\perp^{\mathrm{in}}$, whereas the latter shows $1/N \mathrm{d}\sigma/\mathrm{d}(1-T)$ plotted



(a) (b)

Figure 1: Envelope plots of $p_\perp^{\mathrm{in}}$ (a) and $1-T$ (b). The labels of the y-axes are wrong. It should be $1/N\mathrm{d}\sigma/\mathrm{d}p_\perp^{\mathrm{in}}$ and $1/N\mathrm{d}\sigma/\mathrm{d}(1-T)$ respectively.

against $1-T$. The cross section is σ, $p_\perp^{\mathrm{in}}$ is the in-plane $p_\perp$ with respect to the plane spanned by the major thrust and the minor thrust and T is the thrust. The plots use the complete envelopes of all runs, since the confidence level is set to 100%. One can see that concerning $p_\perp^{\mathrm{in}}$ the generator response envelopes the experimental data very well, whereas the description of $1-T$ causes some problems for $T \to 1/2$. We can account for these problems by assigning a higher weight to the $1-T$ observable, later on. This will cause the Monte Carlo response in the problematic bins to be as close as possible to the actual data. Now, we sample subsets of the total amount of parameter points. Each of these subsets will be used for a distinct tune. By comparing the results of these tunes, we will be able to tell if the amount of parameter points is big enough to secure good statistics for our tune. The PROFESSOR system provides a way to randomly produce these run combinations based on how many runs we want to use per combination.

The next step is the bin-by-bin interpolation of the observables that will be performed for each of the run combinations. Even though a quadratic interpolation is sufficient in almost any case [2], the PROFESSOR system also provides the possibility to perform a cubic interpolation, but for now, we will stick to the quadratic polynomials. Doing both, a quadratic and a cubic interpolation and comparing them afterwards is a good way to validate the parametrisation. Although the observable weights only play a role in the minimisation of the GoF function during the tuning process, it is already necessary to specify them to the PROFESSOR system before the interpolation, since the weight zero might be assigned to some of the observables and the generator responses that correspond to these observables do not have to be parametrised. The assigned weights, which were also used in reference [2], are displayed in Table 1.

After the interpolation, we produce sensitivity plots in order to justify the choice of observables to tune the parameters to. The sensitivity $S_i^{(O)}$ of an observable to a parameter

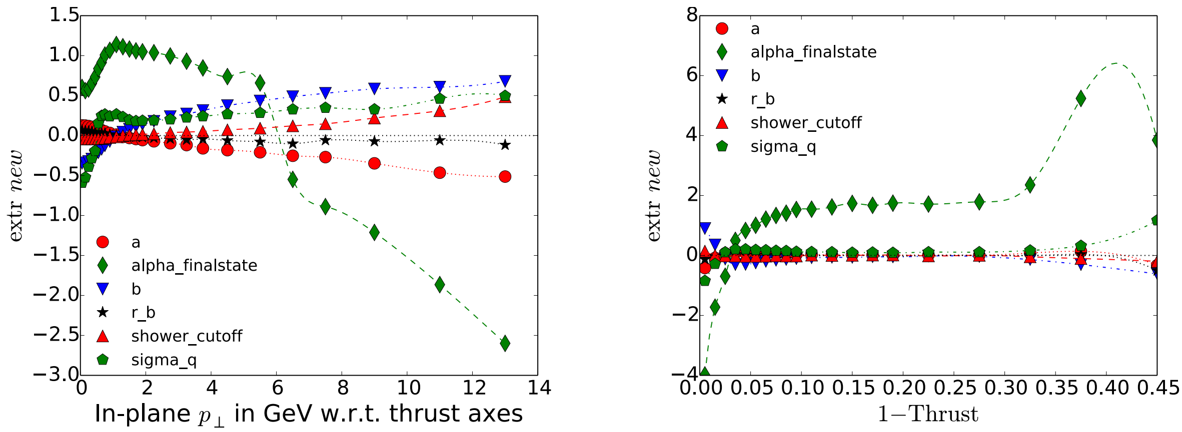
Observable	Weight
$p_\perp^{in}$ w.r.t. thrust axes	1
$p_\perp^{out}$ w.r.t. thrust axes	1
$p_\perp^{in}$ w.r.t. sphericity axes	1
$p_\perp^{out}$ w.r.t. sphericity axes	1
Scaled momentum, $x_p = \|p\|/\|p_{\text{beam}}\|$	1
Log of scaled momentum, $\log(1/x_p)$	1
Mean $p_\perp^{\text{out}}$ vs. x_p	1
Mean $p_\perp$ vs. x_p	1
$1 -$ Thrust	1
Thrust major, M	1
Thrust minor, m	1
Oblateness, $M - m$	1
Sphericity, S	1
Aplanarity, A	1
Planarity, P	1
C parameter	1
D parameter	1
Energy-energy correlation, EEC	1
Mean charged multiplicity	160

Table 1: Observable weights for the Q^2 ordered shower.

p_i can be defined as

$$S_i^{(O)} = \frac{\partial \mathrm{MC}^{(O)}(\boldsymbol{p})}{|\mathrm{MC}^{(O)}(\boldsymbol{p}_t)| + \epsilon w_{\mathrm{MC}}} \frac{|p_{t,i}| + \epsilon w_{p_i}}{\partial p_i}\,, \tag{4}$$

where $\mathrm{MC}^{(O)}(\boldsymbol{p})$ is the generator response to the parameter point $\boldsymbol{p}$ and $\boldsymbol{p}_t$ is a typical set of parameter values. The generator response $\mathrm{MC}^{(O)}(\boldsymbol{p}_t)$ and $\partial\mathrm{MC}^{(O)}(\boldsymbol{p})$ have to be constructed via the interpolation of the generator response, since exact data might not be available for the necessary parameter points. The ϵ-terms have to be introduced to work around the problematic cases $\mathrm{MC}^{(O)}(\boldsymbol{p}_t) = 0$ and $\partial p_i = 0$ and are set to 1% of the parameter range, whereas w_{p_i} is set to 80% of the parameter range and w_{MC} is constructed out of it [8]. The sensitivity plots of the two above-mentioned, exemplary observables are shown in Figure 2. Like all the other observables that were considered in the framework of this tune, our two examples have a high sensitivity with respect to the strong coupling constant $\alpha_{\mathrm{s,f}}$ and lower but considerable sensitivities with respect to the other tuning parameters.
Now, we can use the interpolation results to tune our parameters to the reference data. The minimisation results of the tune are listed in Table 2.
There are several ways to visualize the tuning results. One of them is to look at comparison histograms, similar to the envelope plots listed above, but in this case only the



(a) (b)

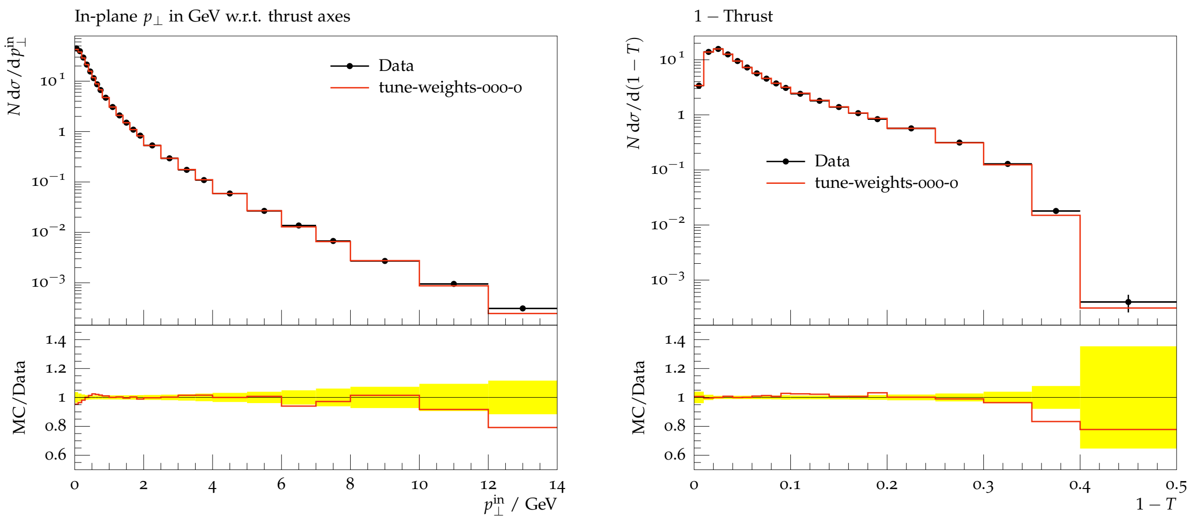
Figure 2: Sensitivity plots of $p_\perp^{\text{in}}$ (a) and $1 - T$ (b).

Parameter	Tune
a	0.2
b	0.4
$\alpha_{\text{s,f}}$	0.143
r_b	0.8
shower cut-off	1.5
σ_q	0.33

Table 2: Results of the tune of the PYTHIA shower with the choice (3) of the longitudinal fragmentation function and without any considerations of photon emission.

best fit result is compared to the reference data. The comparison plots of our two exemplary observables are shown in the Figures 3a and 3b. One can see an acceptable agreement and it is also evident that the method of assigning a higher weight to the problematic observable $1 - T$ was successful as the error in the description of $1 - T$ is now comparable to the one of $p_\perp^{\text{in}}$. The rest of the comparison histograms is listed in the Appendix together with the comparison histogram of the corresponding tune of the WHIZARD shower.

Another way to visualize the tuning results are scatter plots of the optimal parameter values, that belong to the distinct run combinations, mentioned above. There is one scatter plot for each tuning parameter, which shows the GoF measure χ^2 divided by the number of degrees of freedom N_{df} plotted against the optimal value parameter value of each run combination. We hope to find a narrow scattering of the results, because otherwise the plots would indicate, that the result is not sensitive to the tune parameters or there is a systematic error in the procedure. The scatter plots of the considered tune are shown in Figure 4. It also features the scan ranges to make it easier to get a feeling for the scale of the resulting scattering. The plots show that quite a lot of the minimisation results for the parameters a and b are located outside of the scan ranges and the



(a) (b)

Figure 3: Comparison plots of $p_\perp^{in}$ (a) and $1-T$ (b).

scattering of the results of a and the shower cut-off are rather large, which causes the results to be less trustworthy. Because of this, it is advisable, to repeat the tune with adapted scan ranges and more parameter points for the parametrisation.

Further, there is the possibility to perform a scan of the χ^2 function through parameter space along a line that passes through the minimisation result. You can choose the line to be oriented along the steepest or shallowest direction of the covariance matrix of the minimisation result. The PROFESSOR tool would then generate equally distributed parameter points along that line, for which the generator has to be run again. Then, PROFESSOR calculates the χ^2 of these individual runs, which allows us to compare it to our best-fit result. This gives us the possibility to check whether our estimation is close to an actual minimum of the χ^2 function or if it is completely off.

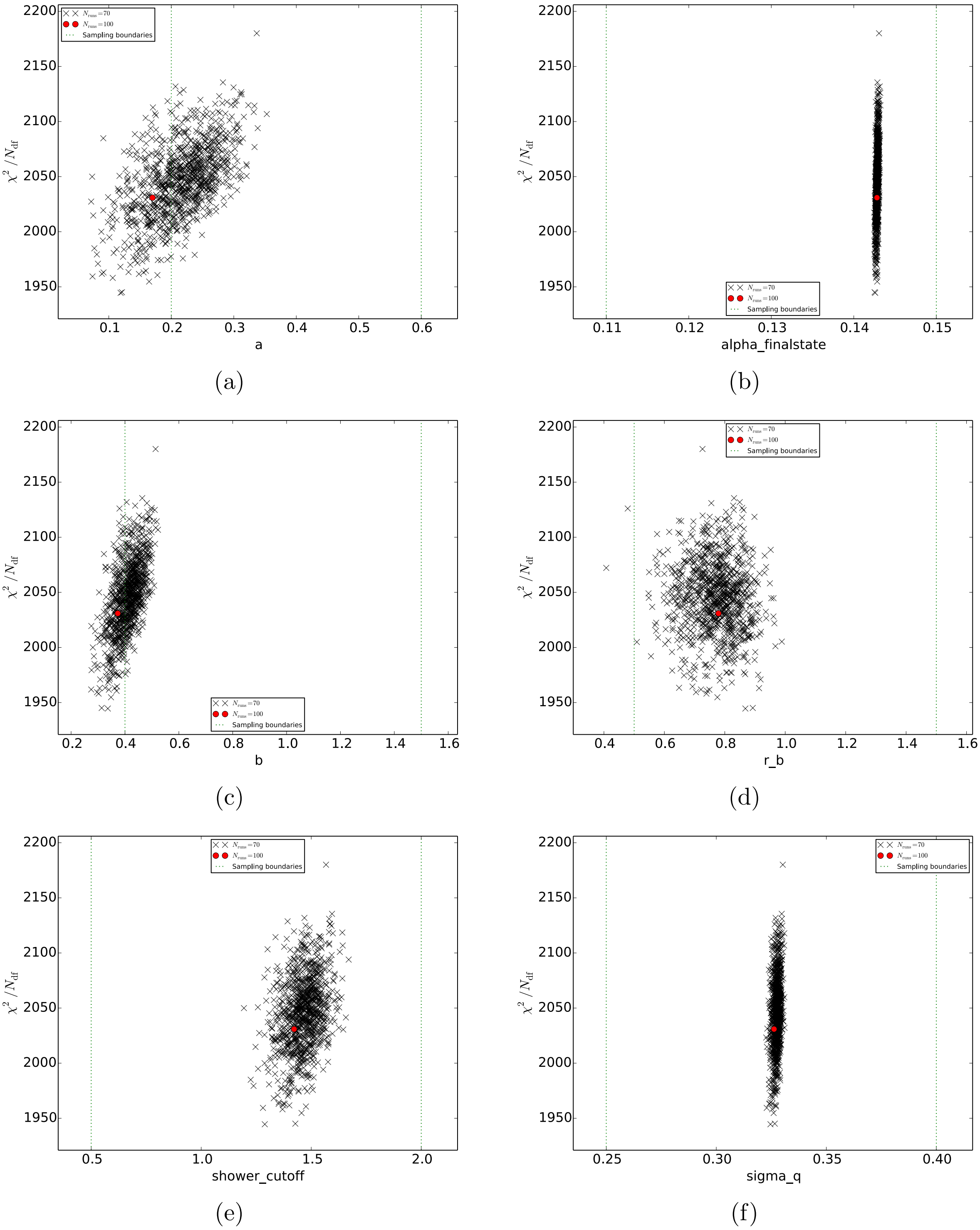


Figure 4: Scatter plots of the parameters a (a), $\alpha_{\rm f,s}$ (b), b (c), r_b (d), the shower cutoff (e) and σ_q (f).

4 Comparisons and summary

Our first aim is to reproduce the tuning results which were achieved in Reference [2] with the PYTHIA shower method, to be sure that everything concerning the tuning process is understood well enough. These results are listed in Table 3.

Parameter	Tune
a	0.5
b	0.6
$\alpha_{\mathrm{s,f}}$	0.143
r_b	0.67
shower cut-off	1.65
σ_q	0.325

Table 3: Tune results of Reference [2].

Considering the uncertainties, that can be observed in Figure 4, the results of the tune of the previous section are fitting well to the reference tune [2]. In the following sections we want to investigate the effects of several settings of our generator on the tune results.

4.1 Different longitudinal fragmentation functions

The tune results for the longitudinal fragmentation functions (1) and (2) look qualitatively similar to the tune that was previously discussed. But because of the large scattering of the tune results of some parameters, and the fact that some of them lie outside of the tune ranges, they have to be repeated with a larger amount of parameter points and adapted sampling ranges to gain more confidence in the tune results.

4.2 QED initial state radiation

Now, we execute a tune based on generator runs, which consider QED initial state radiation. The results are listed in Table 4 and one can see, that they are rather close to the results of Table 2.

Parameter	Tune
a	0.3
b	0.4
$\alpha_{\mathrm{s,f}}$	0.142
r_b	0.7
shower cut-off	1.5
σ_q	0.33

Table 4: Results of the tune of the PYTHIA shower with the choice (3) of the longitudinal fragmentation function and with QED initial state radiation.

The corresponding result scatter plots are qualitatively identical with Figure 4. Therefore, the QED initial state radiation does not seem to have any significant influence on the tune.

4.3 Photon emissions included in the shower evolution, no QED initial state radiation

The next step is to perform a tune without considering QED initial state radiation but adding photon emissions in the shower evolution. The results of this tune are displayed in Table 5.

Parameter	Tune
a	0.3
b	0.4
$\alpha_{\mathrm{s,f}}$	0.142
r_b	0.7
shower cut-off	1.5
σ_q	0.33

Table 5: Results of the tune of the PYTHIA shower with the choice (3) of the longitudinal fragmentation function and considering photon emission, but without QED initial state radiation.

Like before, the tune values are rather close to the ones shown in Table 2. The result scatter plots also are qualitatively similar to Figure 4. We conclude that, within the limits of the uncertainties of the tune, the inclusion of photon emissions in the shower evolution does not affect the results significantly.

4.4 The effect of the number of parameter points on the results

We will now examine a tune that takes into account 1000 instead of 100 parameter points for the interpolation of the observable bins, to study the stability of the tune.
The resulting tune is displayed in table 6.

Parameter	Tune
a	0.2
b	0.4
$\alpha_{\mathrm{s,f}}$	0.142
r_b	0.7
shower cut-off	1.5
σ_q	0.33

Table 6: Results of the tune of the PYTHIA shower with the choice (3) of the longitudinal fragmentation function and considering photon emission, but without QED initial state radiation. About 1000 parameter points were used for the interpolations.

The corresponding scatter plots are shown in Figure 5. It turns out that the scattering is now more narrow and therefore it is advisable to also repeat the other tunes with a larger amount of generator runs. The tune results that are shown in table 6 still agree with the tune of reference [2]. However, we can still see, that some of the tune results are located outside of the sampling ranges.
Based on the results, yielded by the considered tunes, it is not possible to decide, which one fits the reference data the best. The reason for that is, that the scatter plots show a large scattering of the GoF values of the best fit parameters between the different run combinations compared to the difference between the distinct tunes that were performed for this paper.

The reproduction of the tune [2] was successful, but to get more confidence in the tune, the issues that were mentioned, namely the fact that some of the tuning results are located outside of the sampling ranges and that the scattering of the best fits is rather large, remain to be resolved. One approach to solve this problem is to rerun the generator with parameter ranges that are adapted to the results of the previous run. Furthermore, we found that 100 runs do not seem to be sufficient for the tune in some cases, which is why next time, $N \gtrsim 1000$ runs should be used.

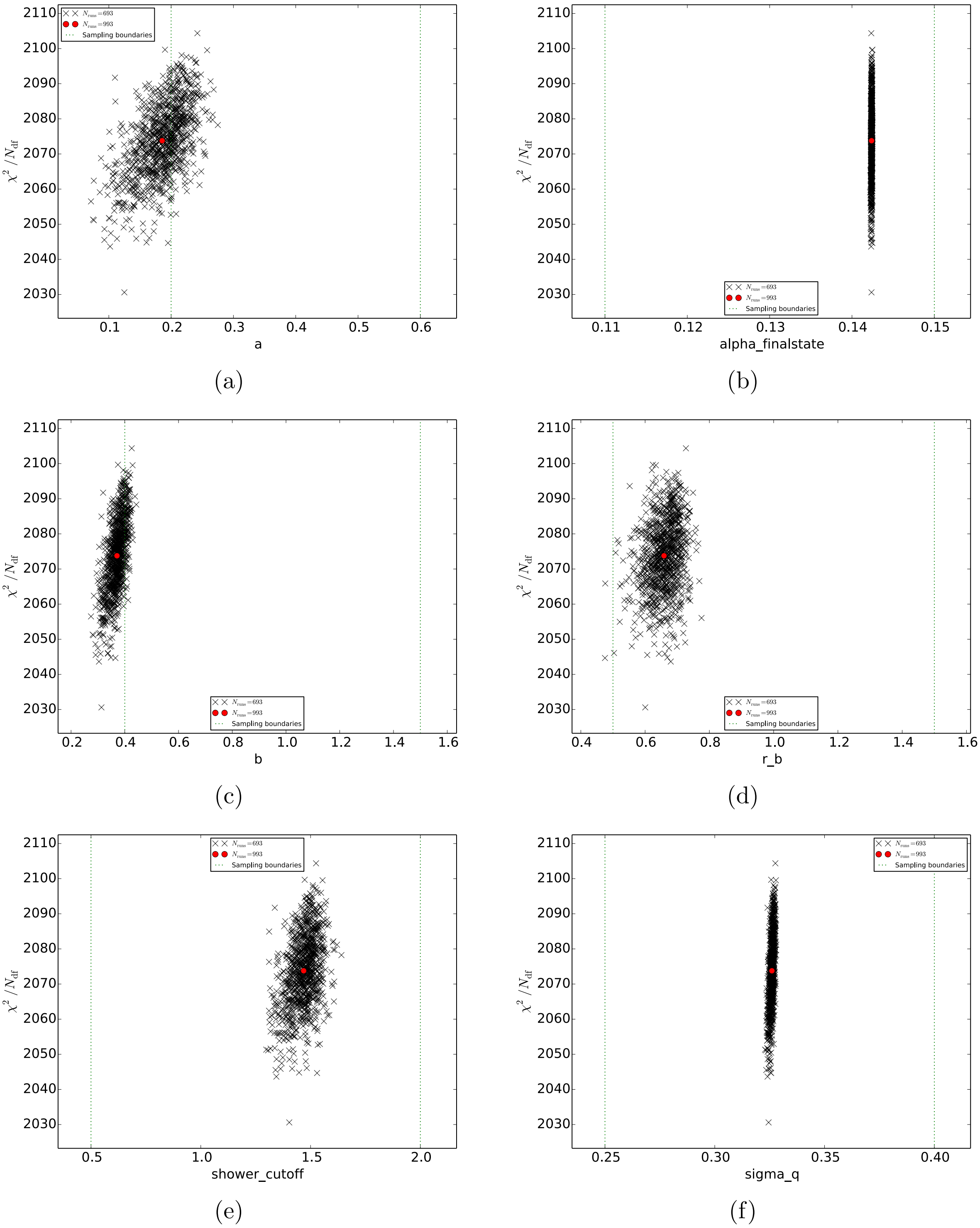


Figure 5: Scatter plots of the parameters a (a), $\alpha_{\mathrm{f,s}}$ (b), b (c), r_b (d), the shower cutoff (e) and σ_q (f).

4.5 Whizard tune

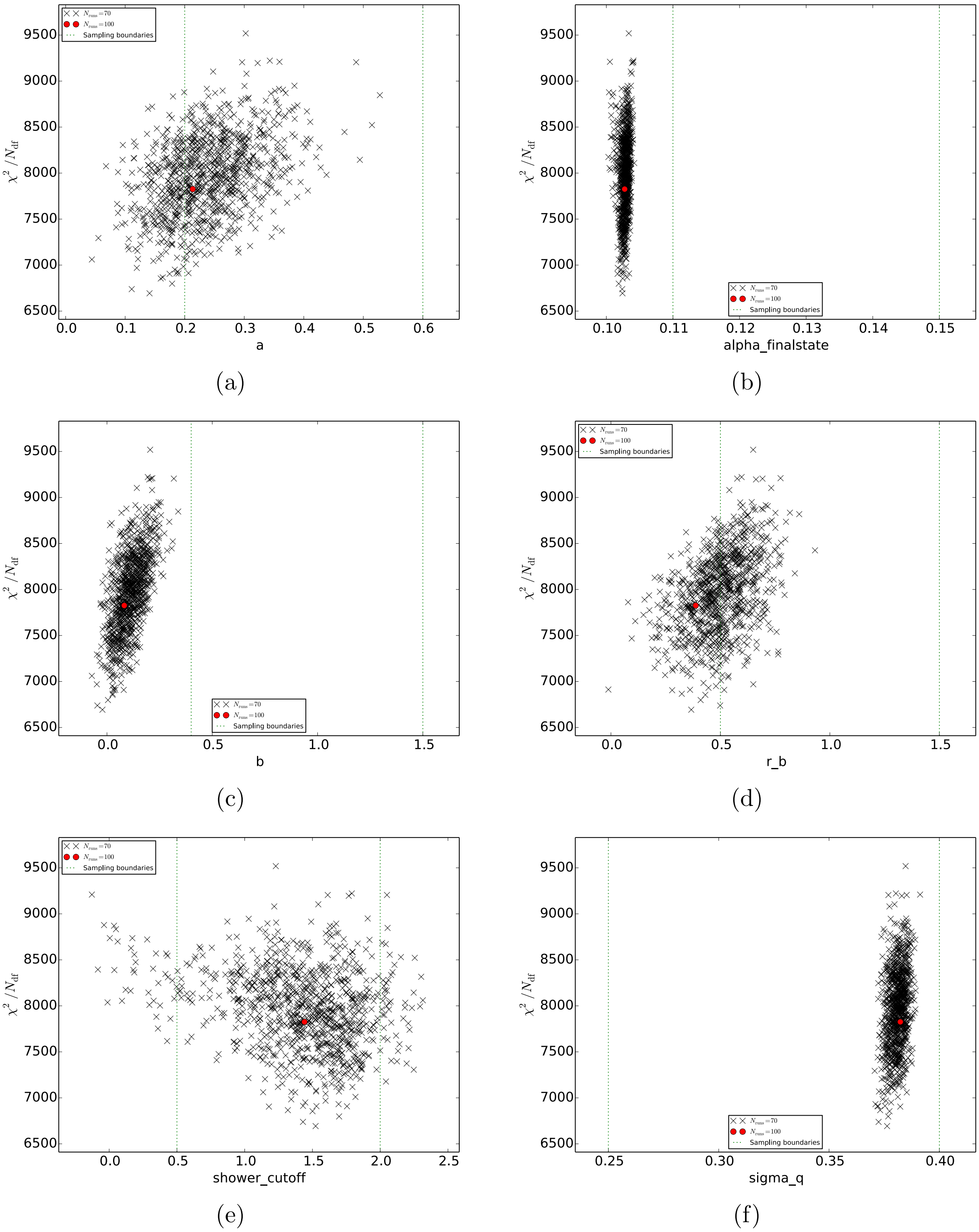
Now, we want to discuss the tunes that were performed for the WHIZARD shower. We only consider the case without addition of photon emission to the shower evolution. The results for the longitudinal fragmentation function (1) are shown in Table 7.

Parameter	Tune
a	0.3
b	0.1
$\alpha_{\mathrm{s,f}}$	0.103
r_b	1
shower cut-off	1.6
σ_q	0.38

Table 7: Results of the tune of the WHIZARD shower with the choice (1) of the longitudinal fragmentation function.

The scatter plots are shown in Figure 6. The scattering is in general even larger than it was observed in the PYTHIA tune and a lot of the results are located outside of the sampling ranges. Therefore, the tune has to be repeated with a larger number of parameter points and adapted sampling ranges.

The results for the longitudinal fragmentation functions (2) and (3) look qualitatively the same. There is a large scattering and some results lie outside of the sampling ranges and therefore the tune is untrustworthy. We conclude that the tunes have to be repeated with more parameter points and adapted sampling ranges.



(a) (b) (c) (d) (e) (f)

Figure 6: Scatter plots of the parameters a (a), $\alpha_{\mathrm{f,s}}$ (b), b (c), r_b (d), the shower cutoff (e) and σ_q (f).

5 Appendix

5.1 Comparison plots

5.1.1 Pythia shower

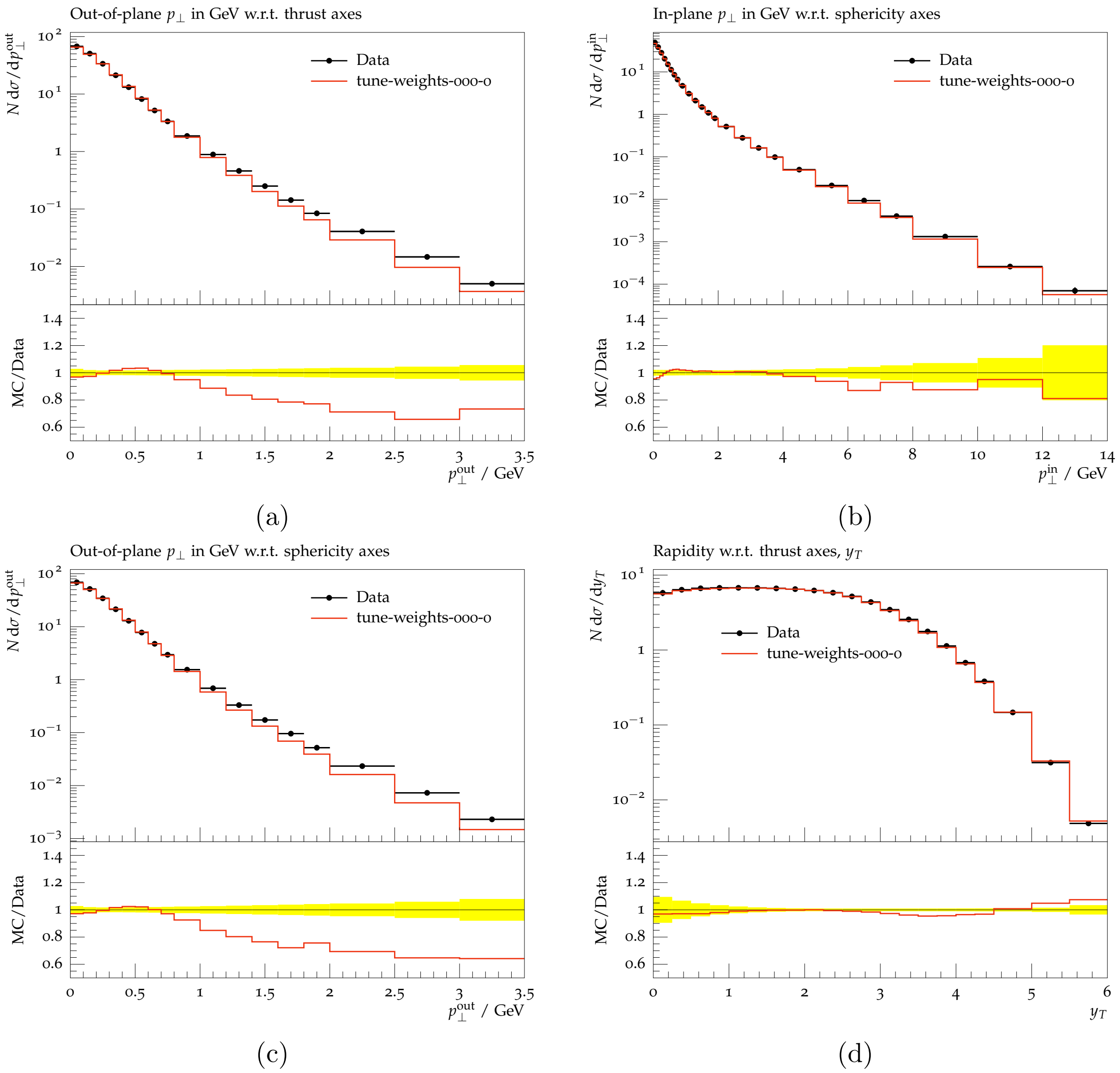
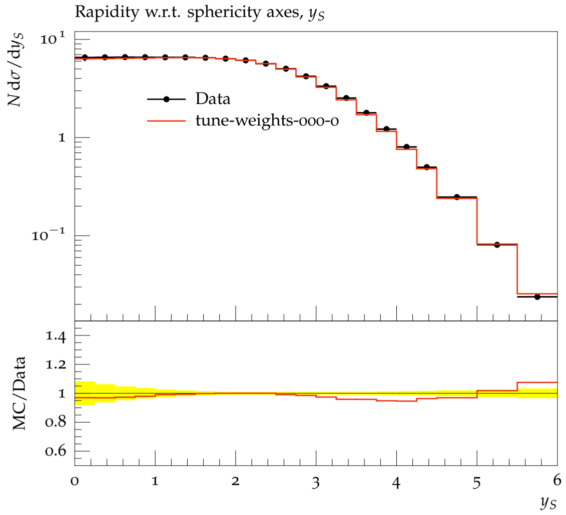
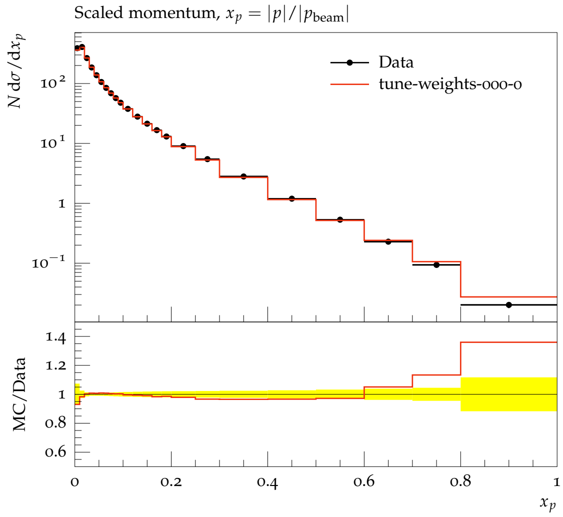


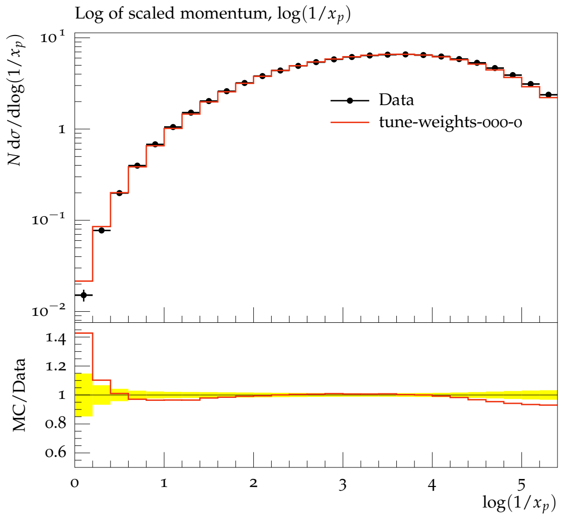
Figure 7: Comparison plots of $p_\perp^{\text{out}}$, w.r.t thrust axes (a), $p_\perp^{\text{in}}$ w.r.t. sphericity axes (b), $p_\perp^{\text{out}}$ w.r.t. sphericity axes (c) and rapidity w.r.t thrust axes (d).



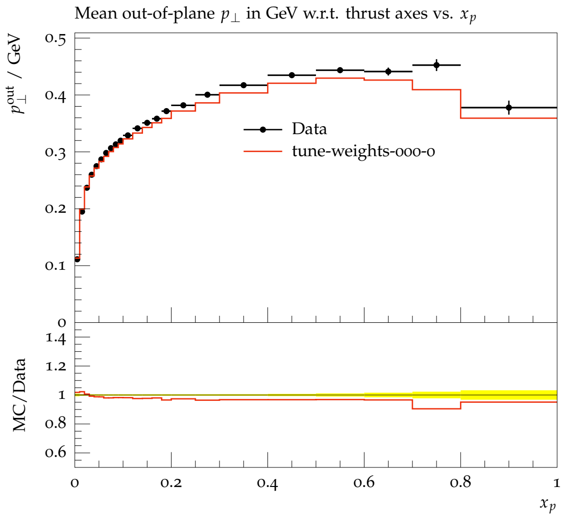
(a)



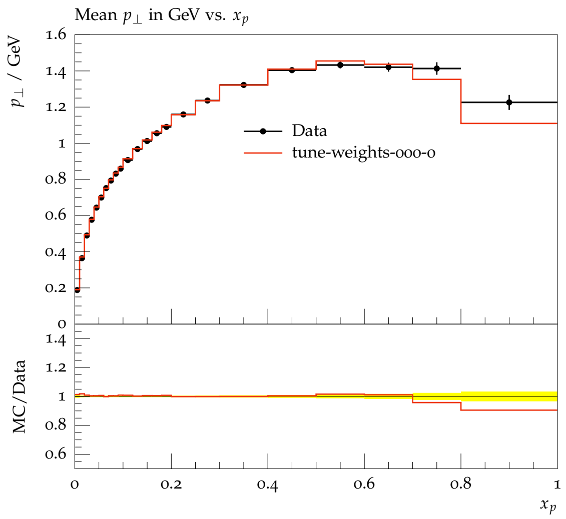
(b)



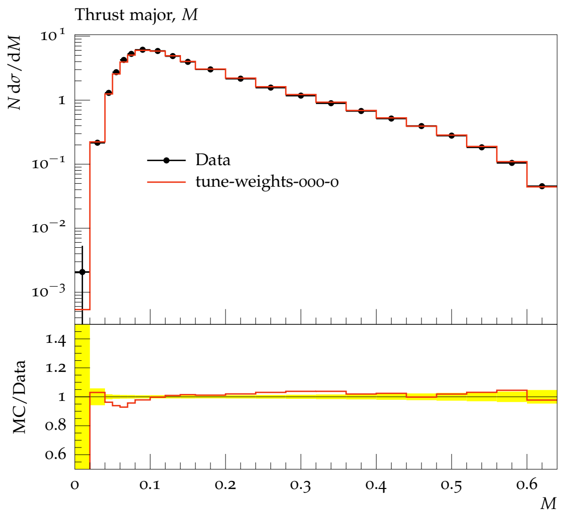
(c)



(d)



(e)



(f)

Figure 8: Comparison plots of rapidity w.r.t. sphericity axes (a), x_p (b), $\log(1/x_p)$ (c), mean $p_\perp^{\mathrm{out}}$, w.r.t thrust axes vs. x_p (d), mean $p_\perp$ vs. x_p (e) and thrust major M (f).

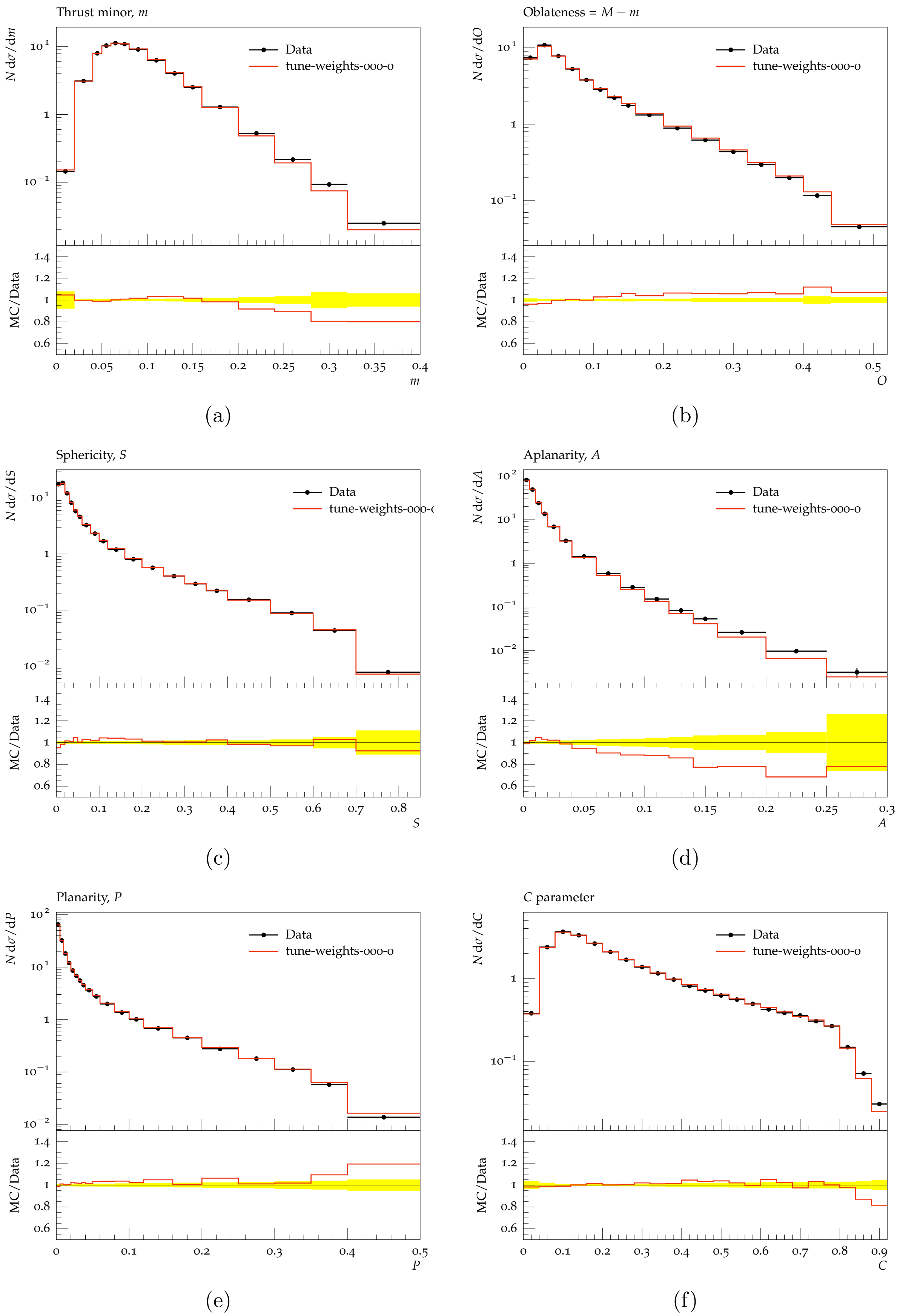


Figure 9: Comparison plots of thrust minor m (a), oblateness $M - m$ (b), sphericity S (c), aplanarity A (d), planarity P (e) and the C parameter (f).

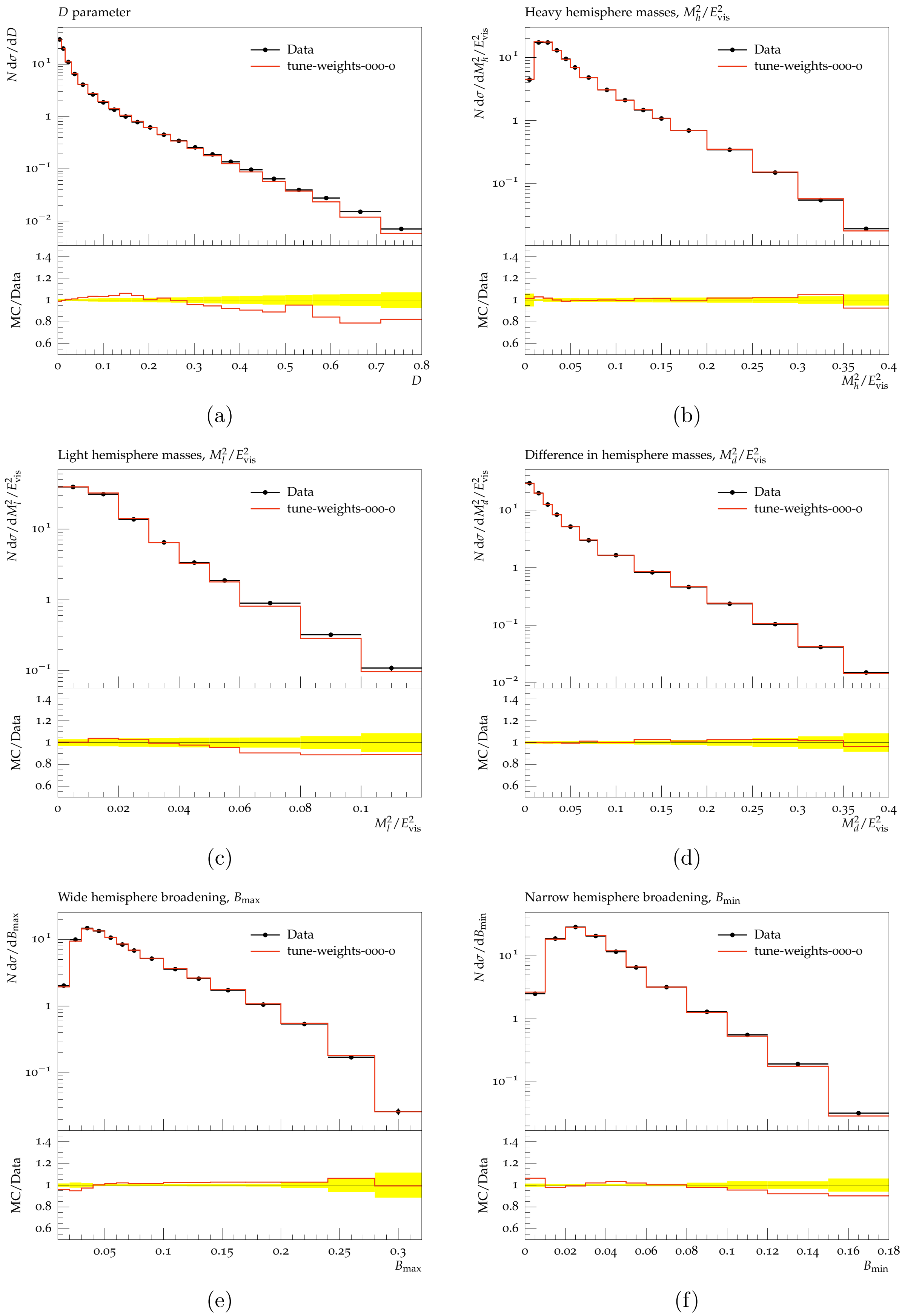


Figure 10: Comparison plots of D parameter (a), heavy hemisphere masses (b), light hemisphere masses (c), difference in hemisphere masses (d), wide hemisphere broadening (e) and narrow hemisphere broadening (f).

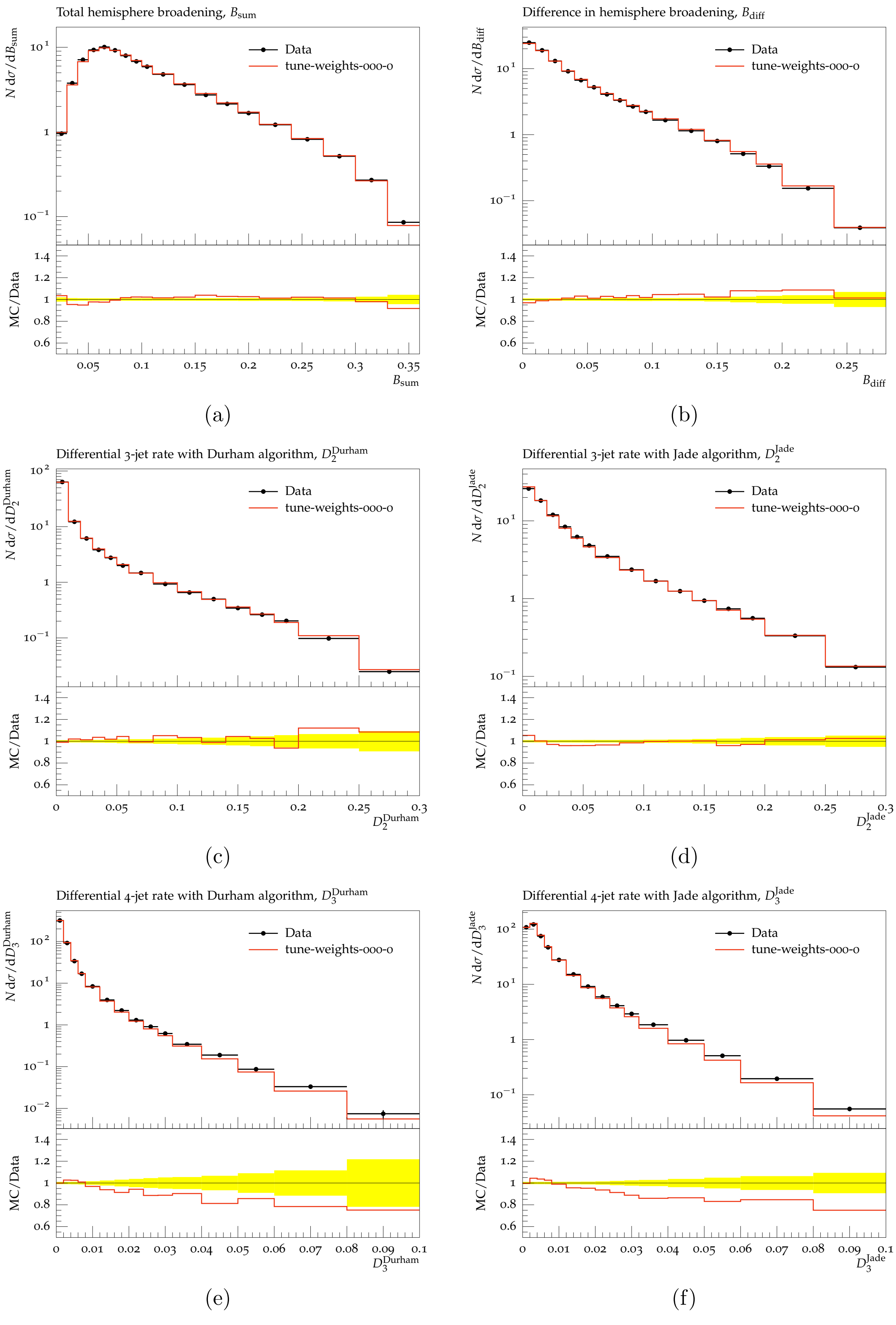


Figure 11: Comparison plots of total (a) and difference in hemisphere broadening (b), differential 3-jet rate with Durham (c) and with Jade algorithm (d), differential 4-jet rate with Durham (e) and with Jade algorithm (f).

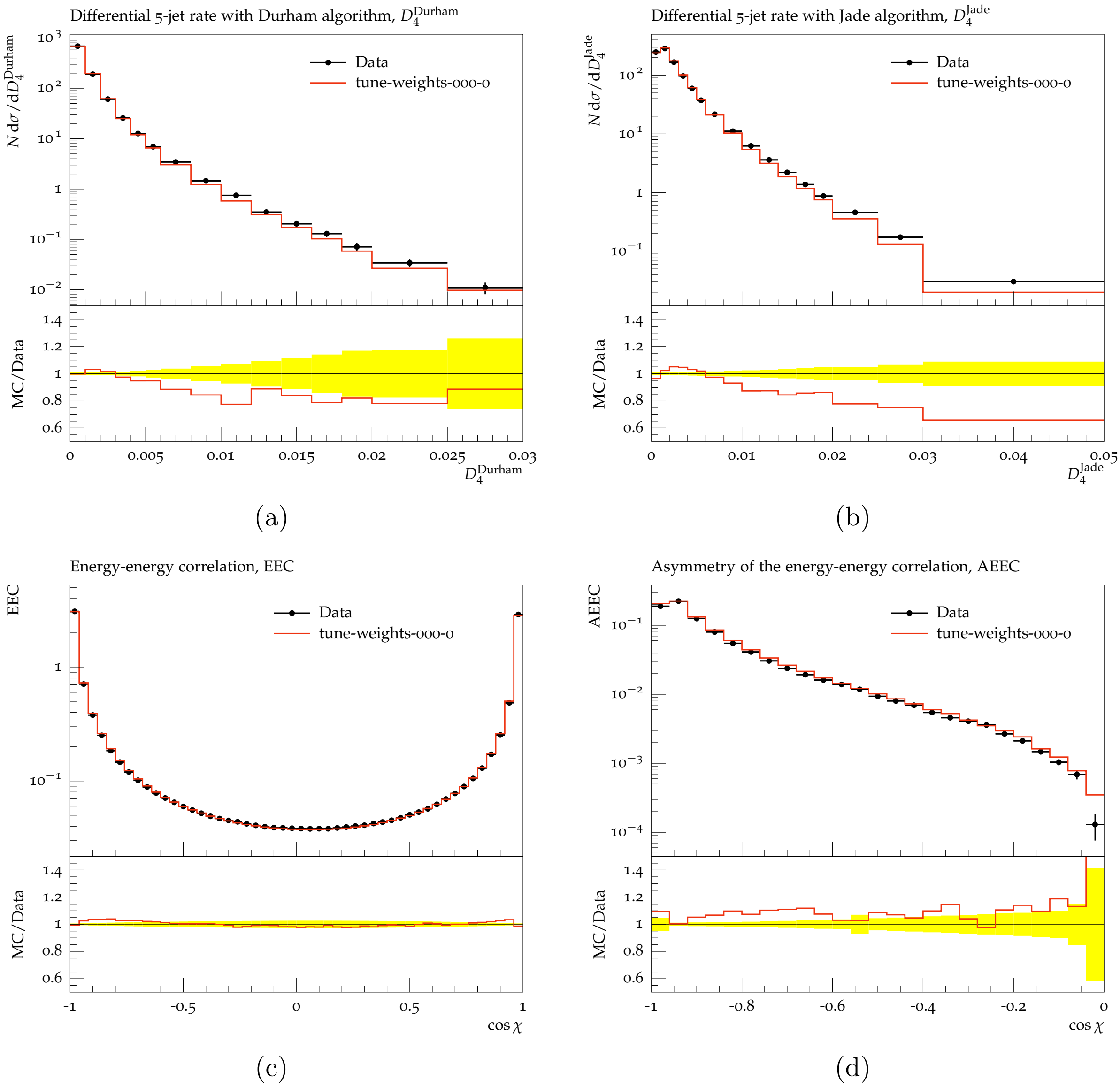


Figure 12: Comparison plots of differential 5-jet rate with Durham algorithm (a), differential 5-jet rate with Jade algorithm (b), energy-energy correlation (c) and asymmetry of the energy-energy correlation (d).

5.1.2 Whizard shower

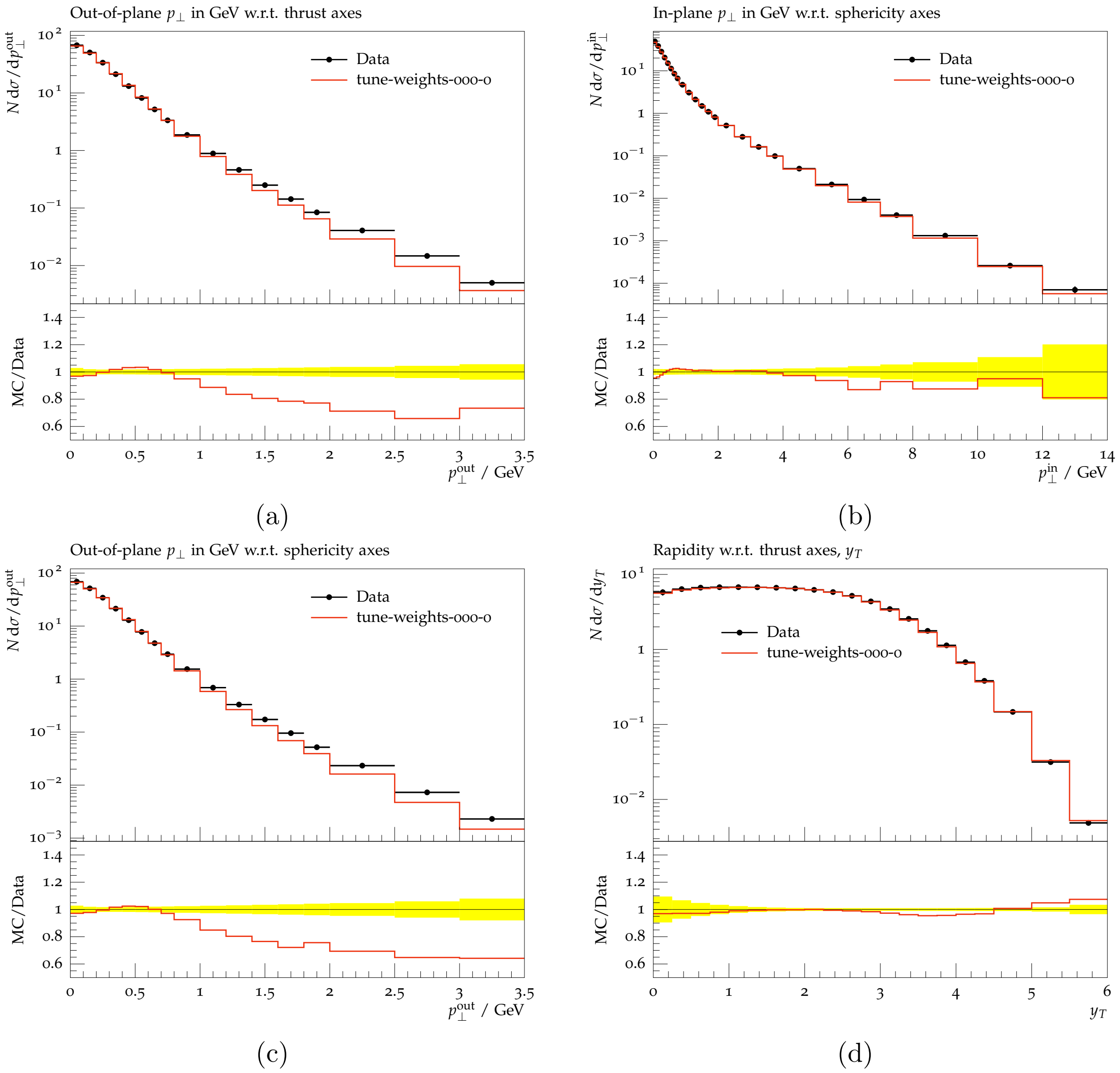


Figure 13: Comparison plots of $p_\perp^{out}$, w.r.t thrust axes (a), $p_\perp^{in}$ w.r.t. sphericity axes (b), $p_\perp^{out}$ w.r.t. sphericity axes (c) and rapidity w.r.t thrust axes (d).

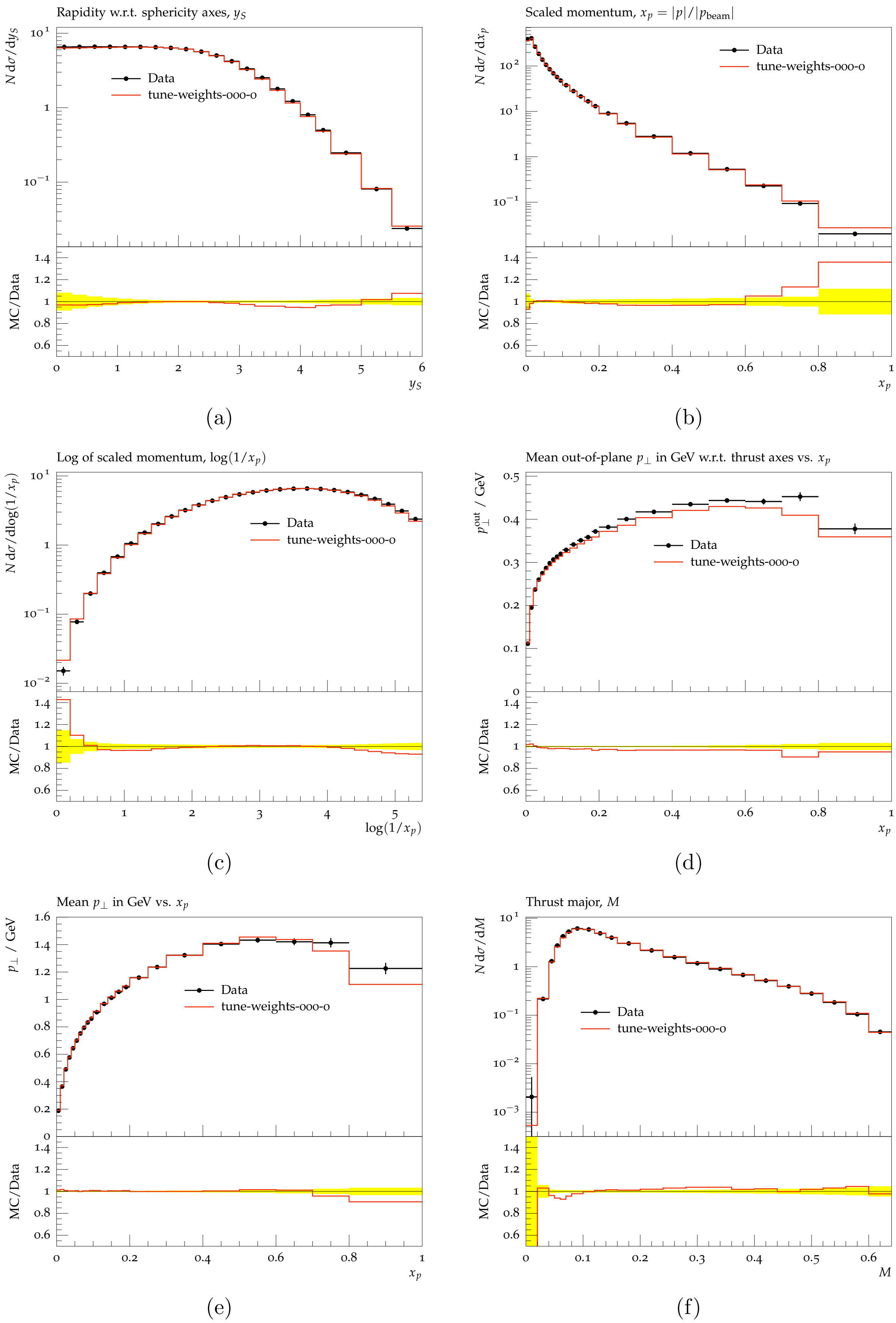


Figure 14: Comparison plots of rapidity w.r.t. sphericity axes (a), x_p (b), $\log(1/x_p)$ (c), mean $p_\perp^{\rm out}$, w.r.t thrust axes vs. x_p (d), mean $p_\perp$ vs. x_p (e) and thrust major M (f).

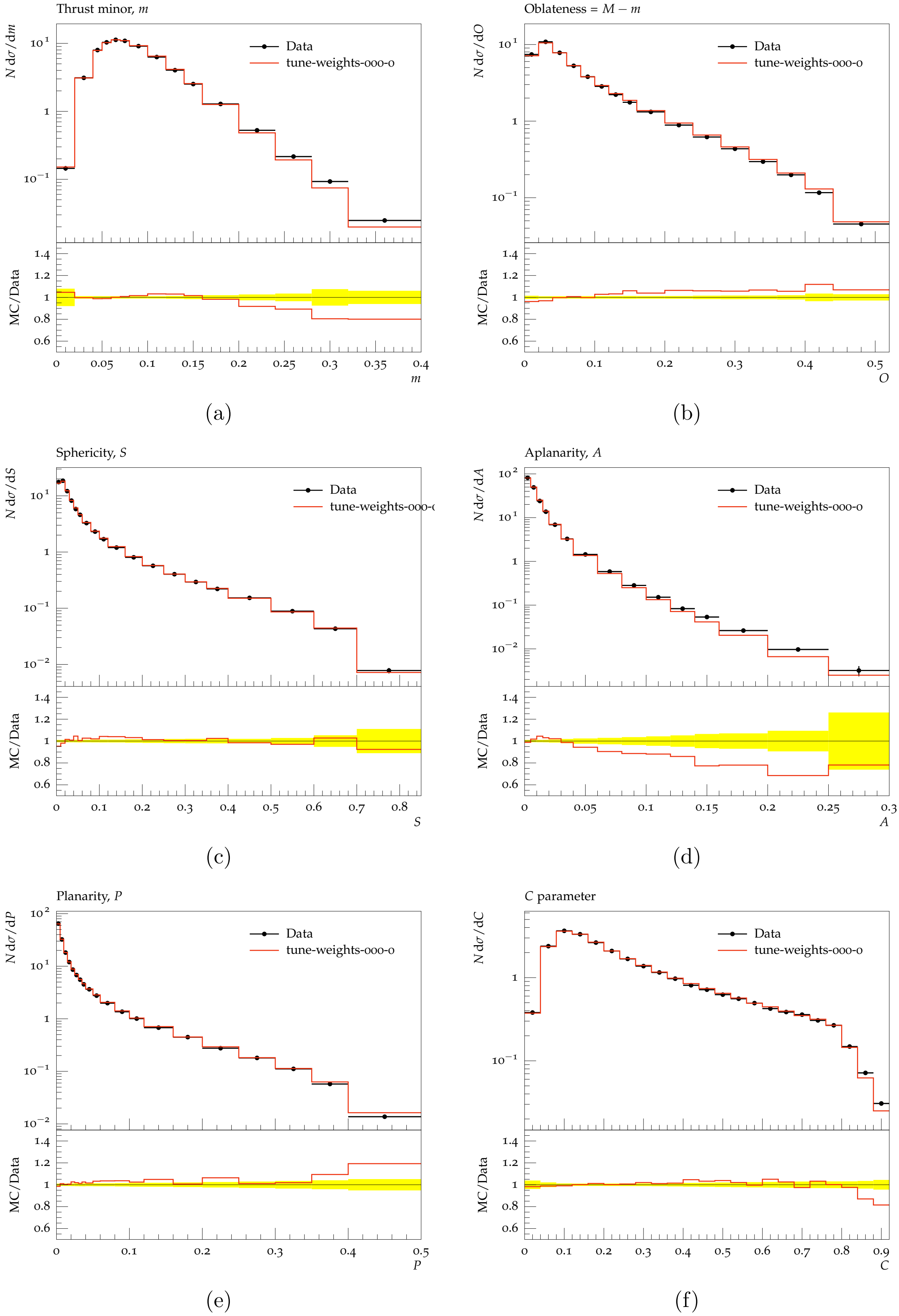


Figure 15: Comparison plots of thrust minor m (a), oblateness $M - m$ (b), sphericity S (c), aplanarity A (d), planarity P (e) and the C parameter (f).

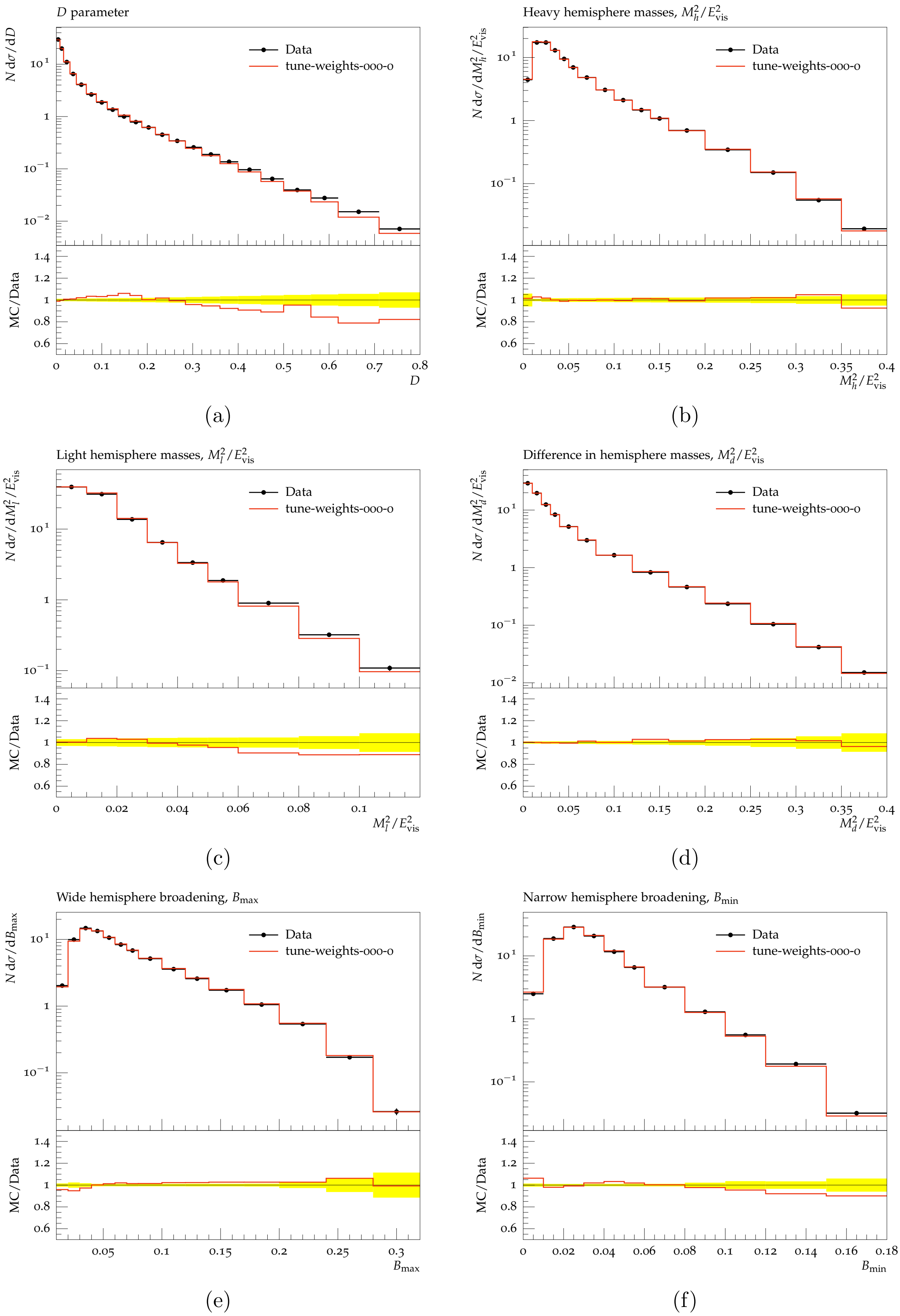
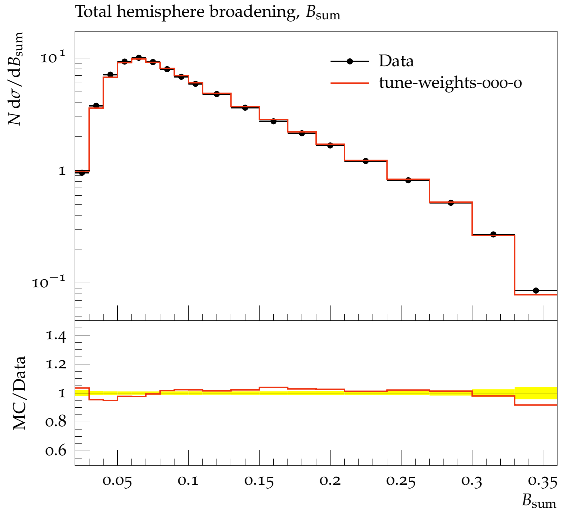
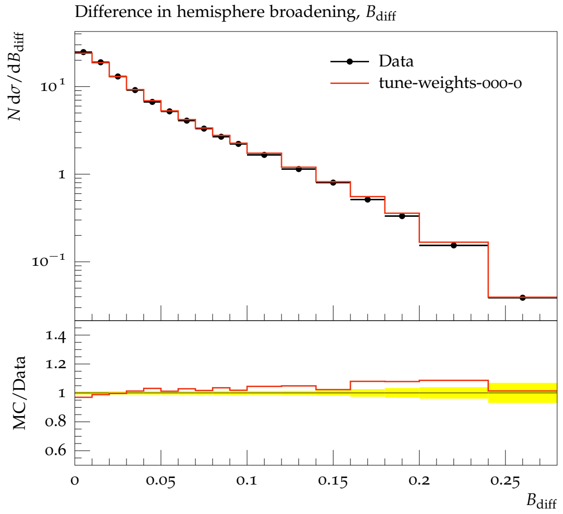


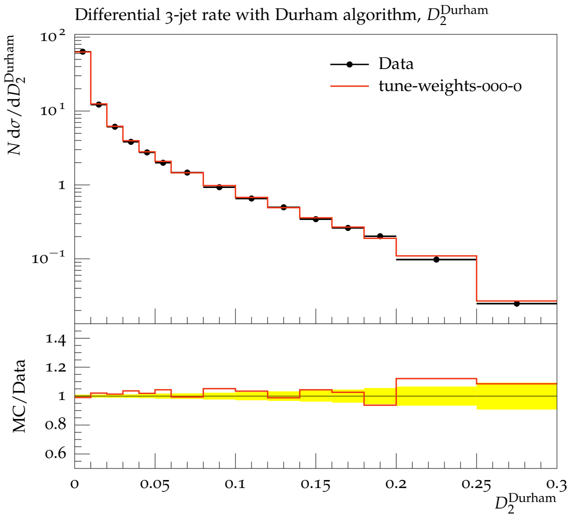
Figure 16: Comparison plots of D parameter (a), heavy hemisphere masses (b), light hemisphere masses (c), difference in hemisphere masses (d), wide hemisphere broadening (e) and narrow hemisphere broadening (f).



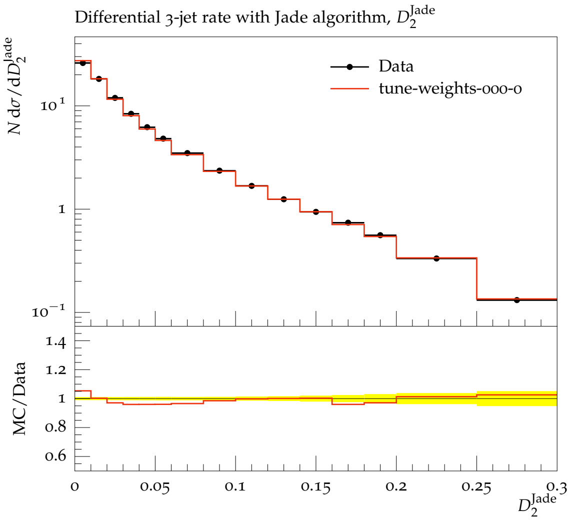
(a)



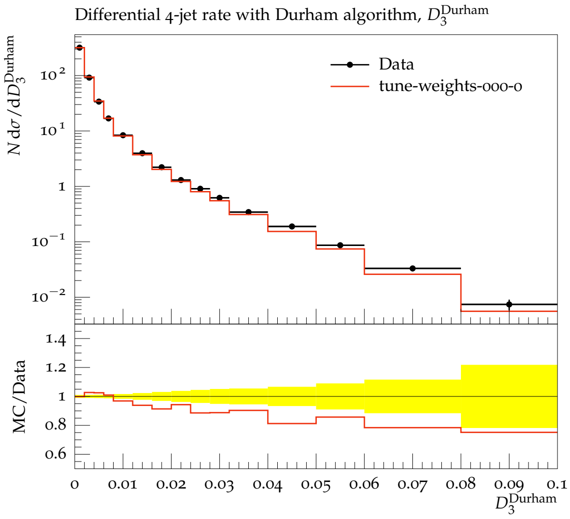
(b)



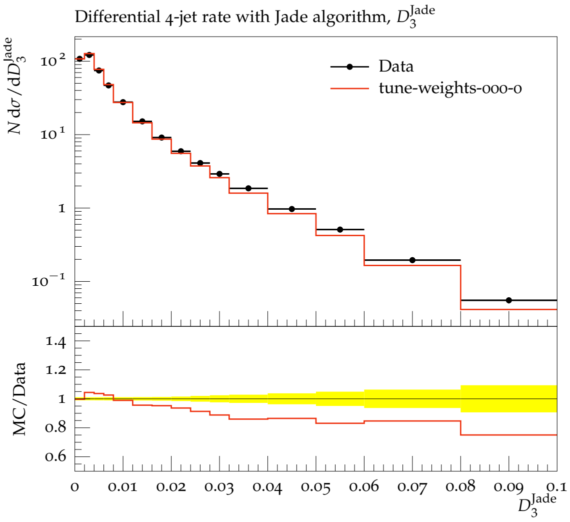
(c)



(d)



(e)



(f)

Figure 17: Comparison plots of total (a) and difference in hemisphere broadening (b), differential 3-jet rate with Durham algorithm (c) and with Jade algorithm (d), differential 4-jet rate with Durham algorithm (e) and with Jade algorithm (f).

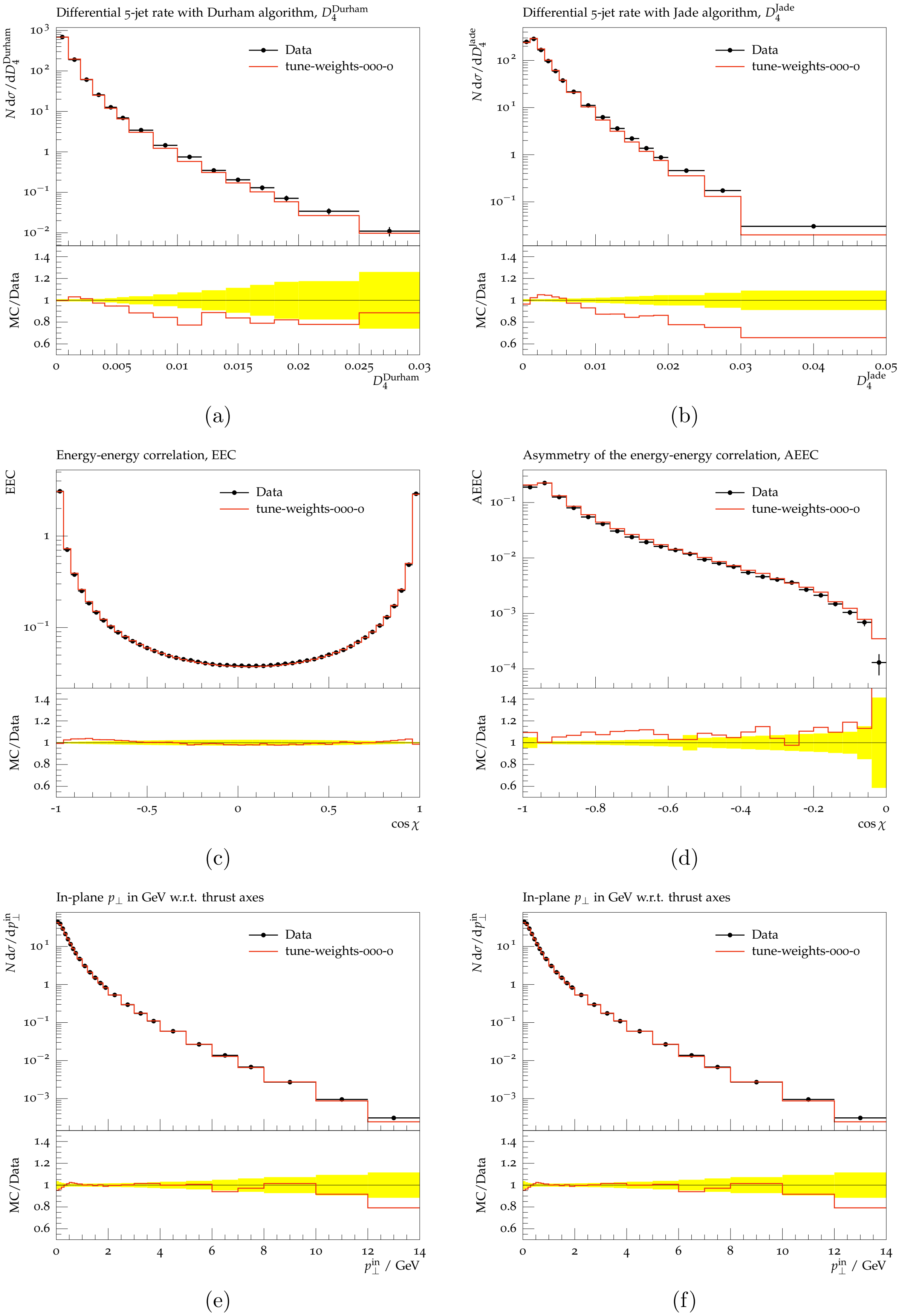
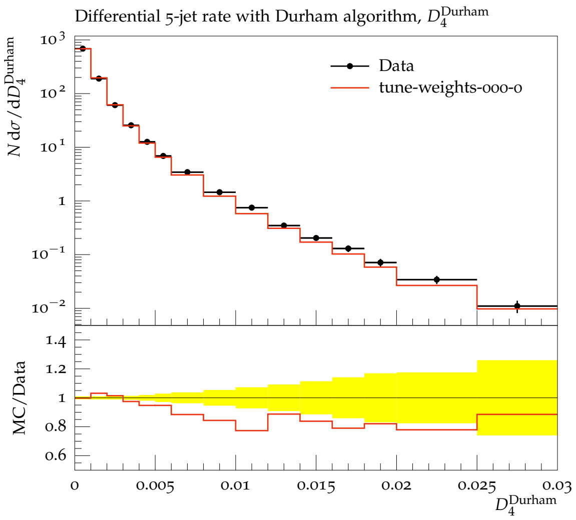
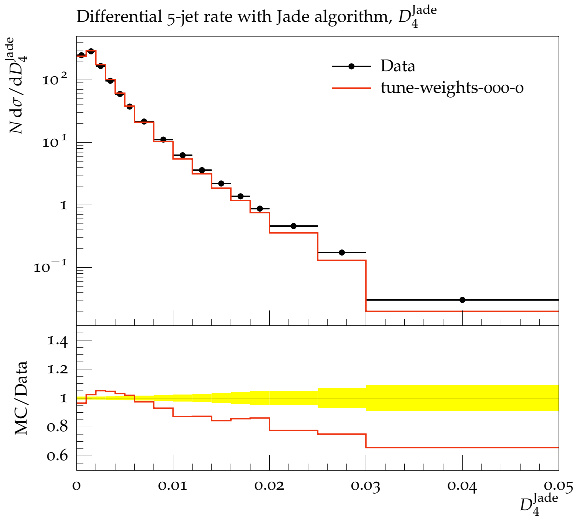


Figure 18: Comparison plots of differential 5-jet rate with Durham algorithm (a), differential 5-jet rate with Jade algorithm (b), energy-energy correlation (c) and asymmetry of the energy-energy correlation.



(a)



(b)

Figure 19: Comparison plots of $p_\perp^{\mathrm{in}}$ (a) and $1 - T$ (b).

References

[1] Monte Carlo Event Generators, arXiv:1304.6677 (2013), *M. H. Seymore et al.*

[2] Systematic event generator tuning for the LHC, Eur. Phys. J. C **65** (2010) 331, *A. Buckley et al.*

[3] Modern particle physics event generation with WHIZARD , J. Phys.: Conf. Ser. **608** (2015) 012063, *J. Reuter et al.*

[4] Regression and the Moore-Penrose pseudoinverse, Academic Press, New York (1972), *A. E. Albert*

[5] General-purpose event generators for LHC physics, Phys. Rept. **504** (2011) 145, *A. Buckley et al.*

[6] RIVET user manual, Comput. Phys. Commun. **184** (2013) 2803, *A. Buckley et al.*

[7] Tuning and test of fragmentation models based on identified particles and precision event shape data, Z. Phys. C **73** (1996) 11, *DELPHI Collaboration*

[8] A study of the sensitivity to the PYTHIA8 parton shower parameters of $t\bar{t}$ production measurements in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS experiment at the LHC, ATL-PHYS-PUB-2015-007 (2015), *ATLAS Collaboration*