Comparing different Jet-Algorithms for photoproduction Monte Carlo data

Clemens Mellein

DESY Hamburg / MPI Munich / RWTH Aachen

Abstract. The following report shows the results of a comparison between the k_t, anti-k_t, Cambridge/Aachen and SISCone jet-algorithms, which were used to analyse direct and resolved dijet photoproduction Monte Carlo data from the Pythia event generator. For the leading jets we demanded cuts of 25, 15 and - if a third jet existed - 10 GeV. For each algorithm, differences of the same observables between reconstructed- and hadron- level were investigated. To find corresponding jets on different levels a matching procedure was introduced, which depends on an angular distance parameter. To optimize this and the jet radius parameter of the different jet-algorithms is one of our main aims.

1. MOTIVATION

Jet-algorithms are one of the most important tools in high energy physics data analysis. They provide an opportunity to reconstruct the partons, which lead to jets in complex events, and thereby at least partially reconstruct the initiating processes. But since the projection of a multiparticle bunch to a simple jet is fundamentally ambiguous - which reflects the variety of existing jet-algorithms - it is important to understand their characteristcs in defining the jets. Therefore we use Monte Carlo data from the Pythia event generator and compare the properties of the defined jets on hadron level with the properties of the jets on reconstructed level.

The HERA experiments aim at measuring various cross sections $ep \rightarrow jets$ to determine the strong coupling α_s and the proton PDFs. In case of photoproduction reactions, which are at least of $\mathcal{O}(\alpha\alpha_s)$, one can also study the PDFs of the photon, which exhibits a hadronic structure in these special reactions. Therefore it is important to identify the best parameter settings for the used jet-algorithms for the special requirements of these experiments. Since in the last years a lot of new jet-algorithms like the anti-k_t and SISCone have emerged, there is obvious interest in testing their characteristics.

(e-mail: Clemens.Mellein@rwth-aachen.de)

*Supervisors: Günter Grindhammer (guenterg@desy.de), Roman Kogler (kogler@desy.de), Aziz Dossanov (aziza@desy.de)

2. BASICS OF ELECTRON-PROTON-SCATTERING

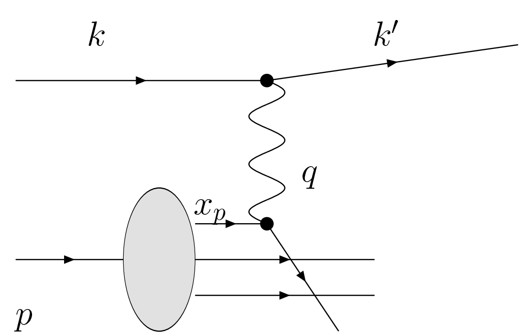

FIG 1. *Electron proton scattering process.*

In electron proton scattering reactions one can observe electromagnetic interactions, mediated by a photon, as well as weak interactions, mediated by the Z_0 (neutral currents) or the $W\pm$ (charged currents). The charged current interactions can easily be separated by cutting on missing energy and momentum, carried away by the emerging neutrino. These scattering reactions (see figure 1) can be described by several sets of kinematical variables. Important variables are the center of mass energy

$$s = (p+k)^2 \approx 4E_eE_p, \tag{2.1}$$

which is usually fixed at collider experiments. E_e is the electron energy, E_p the proton energy. This and all following approximations are valid in the high energy limits, when electron and proton masses can be neglected. The (negative) squared momentum transfer

$$Q^2 = -q^2 = -(k-k')^2 \approx 2E_eE'_e(1-\cos\theta) \tag{2.2}$$

is used to divide the interactions in further classes, like deep inelastic scattering (DIS) with high Q^2. A similar work on jet algorithms for DIS data has been done by B. Lemmer, see (5). In cases of vanishing Q^2 one speaks of *photoproduction* interactions, where a quasi real photon is exchanged. θ is the angle of the scattered electron relative to the beam axis, E'_e the energy of the scattered electron. In reactions $ep \rightarrow eX$ with fixed center-of-mass energy the kinematics can be fully described by two variables,

$$y = \frac{p\cdot q}{p\cdot k} \approx 1 - \frac{E'_e}{E_e}\sin^2\frac{\theta}{2}, \tag{2.3}$$

called the inelasticity of the process, which is in the proton rest frame equal to the energy loss of the scattering electron. The second variable is

$$x_p = \frac{q\cdot a}{q\cdot p}, \tag{2.4}$$

which can be interpreted as the momentum fraction of the scattering parton coming from the proton with four momentum a.

3. PHOTOPRODUCTION

As already mentioned, one speaks of photoproduction in cases of $Q^2 \approx 0$. The photon can either interact directly with one parton from the proton, which is called *direct photoporduction*, or it can fluctuate into a quark-antiquark pair, of which one parton is then interacting with a parton from the proton, this is called *resolved photoproduction*. The longitudinal momentum fraction of the photon-side parton is then defined by

$$x_\gamma = \frac{p\cdot b}{p\cdot q}, \tag{3.1}$$

with the parton momentum b, similar to the proton case (cf. equation 2.4). This means that the photon exhibits a hadronic structure, and one of its partons is participating in the hard interaction.

The usual approach to calculate photoproduction cross sections is to factorize them into one part that is calculable by perturbation theory, like electromagnetic and hard QCD processes, and a non perturbative part, where all QCD processes are parametrized into so called structure functions, which are to be determined experimentally. Therefore we demand at least two hard jets to ensure that the parton interaction is on a hard scale, which is then calculable in pQCD. For the non perturbative processes in photoproduction we get two structure functions $f_p(\alpha_s, x_p, \mu_R, \mu_F)$ and $f_\gamma(\alpha_s, x_\gamma, \mu_R, \mu_F)$, one for the proton structure and the other one for the hadronic structure of the photon. In case of direct photoproduction the latter is not applicable, instead the electromagnetic coupling of the photon with the proton-side parton can be calculated in QED. The two structure functions depend on the strong coupling constant α_s, the proton-side and photon-side momentum fractions respectively, the renormalization scale μ_R and a factorization scale μ_F which defines a boundary for the perturbative and non-perturbative processes. Leading order direct processes are illustrated in figure 2. The cross section for resolved events can be written as:

(3.2)
$$\begin{aligned} d\sigma_{ep}^{res} \quad = \quad & \sum_{i,j=q\bar{q}g} \int_0^1 \int_0^1 dx_p dx_\gamma f_{i,\gamma}(\alpha_s, x_\gamma, \mu_R, \mu_F) \\ & \times f_{j,p}(\alpha_s, x_p, \mu_R, \mu_F) d\sigma_{ij}(\alpha_s, px_p, qx_\gamma, \mu_R) \end{aligned}$$

For direct events:

(3.3)
$$\begin{aligned} d\sigma_{ep}^{dir} \quad = \quad & \sum_{i=q\bar{q}g} \int_0^1 dx_p f_{i,p}(\alpha_s, x_p, \mu_R, \mu_F) \\ & \times d\sigma_i(\alpha_s, px_p, \mu_R) \end{aligned}$$

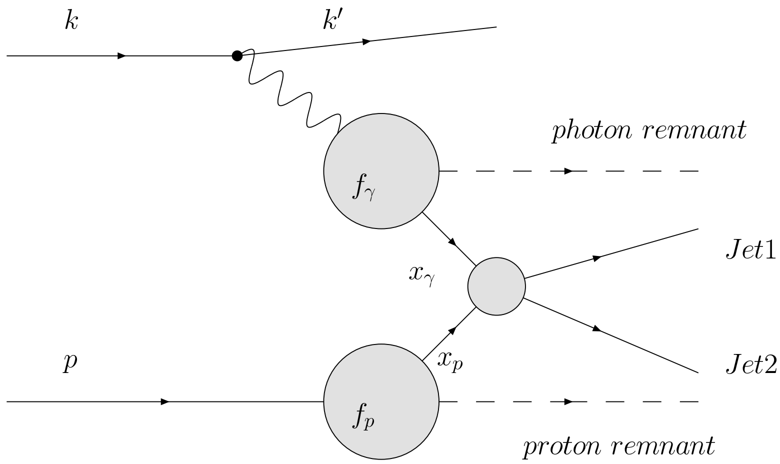

FIG 2. *Schematic view of dijet photoproduction process in ep-scattering. The small grey circle depicts the hard (short scale) parton-parton interaction.*

The leading order processes of direct dijet photoproduction are QCD compton scattering $\gamma q \to gq$ and the so called photon-gluon-fusion $\gamma g \to q\bar{q}$, respectively, which are of $\mathcal{O}(\alpha\alpha_s)$ (cf. figure 3). There are various LO diagrams for the hard parton-parton interaction of the resolved photoproduction (generically depicted with a small grey circle in figure 2), which are of $\mathcal{O}(\alpha\alpha_s^2)$, e.g. processes like $q\bar{q} \to q\bar{q}$, $qq \to qq$ and $qg \to qg$. For these and further NLO diagrams, virtual and real corrections, see (3).

4. JET-ALGORITHMS

For all types of jet-algorithms, one requires infrared, collinear safety and factorizability. Additionally, small renormalization scale dependence and small hadronization corrections are derivable. This guarantees that the defined jets do not lead to divergent observables. Finally, when extracting the PDFs and strong coupling α_s from jet measurements, they should give the same results, independent of the jet-algorithm which was used. On detector level one considers all objects with associated 4-vectors, which have been reconstructed using information from the tracking detectors and calorimeters (reconstructed level). This is the only known information after conducting an experiment. But what one wants to know, the "physical truth", are the 4-momenta of the jets at hadron level, which are expected to provide a close correspondence to the jets at parton level.

In general, one can discriminate between two widespread classes of jet-algorithms, the sequential recombination algorithms (also called clustering algorithms), like k_t, anti-k_t and Cambridge/Aachen and the cone-type algorithms, such as SISCone. Both classes of jet-algorithms are sensitive to different types of non-perturbative QCD corrections, cf. (1) and (2).

For the clustering algorithms some kind of angular distance d_{ij} between two particles or protojets and a distance d_{iB} between a particle (or protojet) and the beam is introduced. The latter distance is used as a stopping criterion of the clustering process, as described further down. The algorithm then proceeds by identifying the minimal d_{ij} and combines i and j, as long as d_{ij} is smaller than d_{iB}. If it is no longer smaller, i is defined as a jet. The merging is done due to the p_t-scheme:

$$\begin{aligned} p_{t,ij} &= p_{t,i} + p_{t,j}, \\ \eta_{ij} &= \frac{p_{t,i}\eta_i + p_{t,j}\eta_j}{p_{t,ij}}, \\ \phi_{ij} &= \frac{p_{t,i}\phi_i + p_{t,j}\phi_j}{p_{t,ij}}. \end{aligned} \tag{4.1}$$

This merging scheme results in massless jets. The jets are ordered ascendently in p_t. The above described procedure is then repeated until no single entity is left. For the three clustering algorithms considered, the distances are defined by:

$$d_{ij} = \min\left(p_{t,i}^{2p}, p_{t,j}^{2p}\right)\frac{\Delta_{ij}^2}{R^2} \tag{4.2}$$

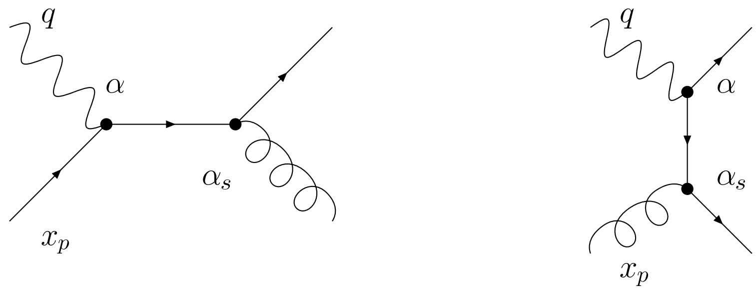

FIG 3. *Leading order diagrams of direct photoproduction with timelike or spacelike quark propagator.*

$$d_{iB} = p_{t,i}^{2p} \tag{4.3}$$

This definition for d_{ij} is composed of the transverse momenta of two particles or protojets, each to the power of $2p$, and an angular distance term, which is governed by the radius parameter R, which is usually of $\mathcal{O}(1)$. The distance Δ_{ij} is defined by:

$$\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2, \tag{4.4}$$

the so called pseudorapidity η is related to the angle θ of a particle (or protojet) via

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{4.5}$$

It is in the high energy limit with negligible particle masses numerically close to the rapidity y defined in special relativity as

$$y = \frac{1}{2}\ln\frac{E + P_z}{E - P_z}. \tag{4.6}$$

It will be one of our main aims to find a good value for the parameter R for each jet-algorithm. Considering the parameter p, the case $p = 1$ corresponds to the k_t algorithm, $p = 0$ to the Cambridge/Aachen algorithm and $p = -1$ to the anti-k_t algorithm. The sign of the parameter p affects the merging of the particles or protojets, considering a hard particle and a soft particle, for positive p the d_{ij} is dominated by the momentum of the soft paritcle, in contrast, in the case of negative p d_{ij} is dominated by the momentum of the hard particle (or protojet). For the special case p=0 only the angular distance of the particles or protojets is relevant. A detailed description of the clustering behaviour of the anti-K_T algorithm you can find in (1).

The general idea of cone-type algorithms is to define jets as a cone around certain directions with high energy flows. Also in this case a geometrical parameter R appears, in this case it is the radius of "trial" cones, which are put around certain seed-particles in the considered event. Then for each seed the sum of the four momenta of all particles in the cone is calculated, which yields a new direction for the trial cone. If the direction of a trial cone no longer changes after a certain amount of iteration steps, the cone is reffered to as a "stable cone". Obviously the sum of the for momenta of all particles in a stable cone then has to coincide with the cone axis. Unfortunately, for the algorithms using seeds problems like infrared and collinear unsafety occur. A way out of these problems is provided by the recent seedless cone algorithm, the SISCone algorithm. For more details about this algorithm see (2).

5. THE MONTE CARLO DATA

The Monte Carlo data used for this analysis is created with the Pythia 6.4 event generator (6) and contains 5 million direct and resolved photoproduction events. Multiple interactions between the resolved photon and the proton are included in this data set. The data are fully simulated and reconstructed, so that a comparison between jets on reconstructed and hadron level is possible.

6. PHASE SPACE DEFINITIONS

For this work we consider photoproduction events with at least two hard jets, their transverse momentum is demanded to be greater than 25 GeV and 15 GeV, respectively. These cuts ensure that the parton parton interaction is hard enough, which is required to trigger on photoprduction events with high efficiency and makes the cross sections calculable with the means of pQCD. If a third jet occurs, also here a cut of 10 GeV is applied. These p_t cuts are equal on hadron and reconstructed level, see table 1 and 2. The η cuts apply for all of the first three jets[1] and are also the same on both levels. We demand η to be in the range of $-0.5 < \eta < 2.75$, which is due to the acceptance of the detector. To remove DIS events we require that there is no electron candidate found during the reconstruction, which restricts the photon virtuality to Q^2 <4 GeV, according to the detector acceptance of HERA II (cf. the relation between Q^2 and the angle θ of the scattered electron in equation 2.2). The same cut on Q^2 is explicitly applied on hadron level. The cuts on the jet masses on rec-level have the purpose to remove any fake jet events, which occur when an electron is reconstructed as a jet after the detector simulation. Real jets have usually masses greater than 2 GeV. Charged currents are excluded with a cut on missing p_t, which has to be below 20 GeV. Since $E - P_z$ is related to y_{JB} via $y_{JB} = \frac{E-P_z}{2E_e}$ the cuts on these quantities correspond to each other on the rec- and had-level.

A cut on the z-vertex position and the non-ep-background finding algorithms are usually used to remove cosmics, but is not necessary in the analysis of Monte Carlo data, since cosmics are not part of the simulation. Anyway, for the sake of completeness we take them in here.

For more details on photoproduction cuts and the above mentioned non-ep-background finders see (4).

TABLE 1
Phase space definition. Cuts on reconstructed level.

parameter	*Cut*
z-vertex pos.	-35 cm< z <35 cm
$p_{t,miss}$	<20 GeV
$E - P_z$	5.52 GeV< $E - P_z$ <49.68 GeV
no scattered electron	
$p_{t,1}$	>25 GeV
$p_{t,2}$	>15 GeV
$p_{t,3}$	>10 GeV
$\eta_{all\ jets}$	$-0.5 < \eta < 2.75$
$M_{jet,1}$	>2 GeV
$M_{jet,2}$	>2 GeV

TABLE 2
Phase space definition. Cuts on hadron level.

parameter	*Cut*
Q^2	<4 GeV^2
y_{JB}	$0.1 < y < 0.9$
$p_{t,1}$	>25 GeV
$p_{t,2}$	>15 GeV
$p_{t,3}$	>10 GeV
$\eta_{all\ jets}$	$-0.5 < \eta < 2.75$

7. RESULTS

7.1 Jet Matching

To compare jets at reconstructed and hadron level, we had to introduce some matching procedure, which yields a pair of jets that is reasonable to compare. To this end, we calculate the distance $\Delta R^2 = (\Delta\eta)^2 + (\Delta\phi)^2$ for the three hardest had-jets to each of the five hardest rec-jets, the minimal distance was taken as matching criterion. We also introduced a boundary parameter ΔR[2], which excludes all those rec-jets as a matching partner that have a greater distance to a certain had-jet than the given value. We calculated the matching efficiency, the ratio of the number of all matched jets to all matching trials, for different values of ΔR to find a parameter region, where the matching effienciency is better than 90% and the sensititvity of the efficiency with respect to the radius parameter of the jet-algorithms is low. A too large ΔR would lead to accidental mismatchings, so that a comparison between the different levels would not be very accurate. A too small ΔR would lead to a small matching efficiency and low statistics in all further analysis. We chose $\Delta R = 0.1$ as a compromise.

Figures 9,10 and 11 show the matching efficiency for a fixed value of $\Delta R = 0.1$ plotted versus p_t, η and ϕ. In the ϕ plot no structure should be recognizable, which is obviously confirmed by figure 11. But the latter also reflects very clearly the stronger dependency of the k_t algorithm on R_0 compared to the anti-k_t. The other two algorithms do not show any peculiar behaviour in terms of the jet matching. In figure 10 one can recognize that the efficiency

[1]The labels 1st, 2nd and 3rd are due to the transverse momentum ordering of the jets.

[2]Dont confuse it with the radius parameter R_0 of the jet-algorithms.

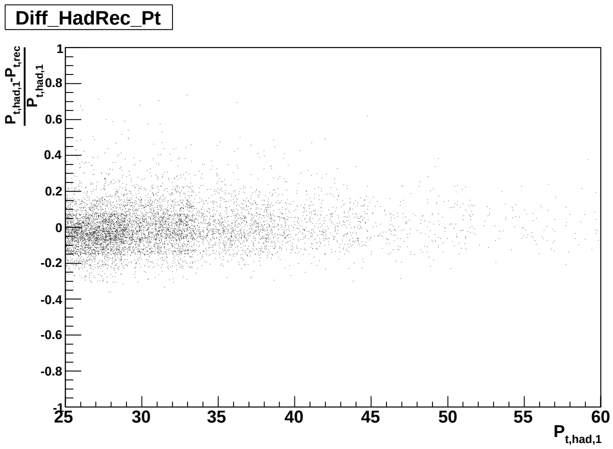

FIG 4. *Shown is, as an example, the 2dim distribution of the difference in p_t between the 1st had-jet and its matching jet on rec-level on the y-axis versus the 1st had-jets p_t on the x-axis. For each x-Bin we calculate the mean value and the sigma of this distribution and fill it into separated histograms (shown later) to get a comprehensive view of the differences between the hadron level and reconstructed level jet definitions for each algorithm. Used algorithm: k_t with $R_0 = 1.0$*

falls off for bigger values of η. Due to several effects like dead detector material, less efficient tracking and scattering of particles coming from the colimator getting into the forward calorimeter, there may be some weird mergings of jets or particles, whereby the matching efficiency finally decreases. In 9 one can see that the efficiency is increasing with increasing p_t, which is due to smaller average angular spread of particles and thereby better defined jets at high p_t. This is also reflected in figures 6, 7 and 8, where the matching efficiency is plotted for each of the first three jets versus ΔR. The efficiency considerably decreases for the 2nd and 3rd had-jet.

7.2 Comparison between reconstructed and hadron level jet quantities

To compare reconstructed and hadron level we use the above described matching procedure starting from jets at hadron level. For example, for every 1st had-jet the difference in p_t, η and ϕ between this jet and its matching jet on rec-level is recorded in a Root TH2D (see fig. 4) versus the had-jet's p_t, η, ϕ and versus the invariant mass of the first two had-jets M_{12}^2. Afterwards, we calculate for each x-Bin the mean value and sigma of the particular difference distribution and fill them into new histograms. This allows to provide a comprehensive view on the differences between the hadron level and reconstructed level jet definitions for each algorithm and for different radius parameters.

- **p_t histograms (figures in appendix B):** in terms of the p_t difference between rec- and had-level, bigger radii seem to improve the matching between the two levels, at least for high p_t of the first had-jet. Here, the SIScone shows the worst performance, k_t and anti-k_t are preferred as they show the smallest differences between rec- and had-level p_t. In terms of the η difference, Cambridge/Aachen and anti-k_t are a little more sensitive to changes in R_0 than the other two algorithms. All algorithms give a positive offset, which gets smaller with decreasing R_0, except for SISCone. Also noticeable is a negative offset in the ϕ difference, which one would not expect. It means that the jet-algorithms systematically find the jets on had-level at smaller ϕ values. All algorithms with all radii show this anomaly in a similar way, one will find it also in all following plots with a similar order of magnitude (≈ -0.005). So far, there is no explanation for it.
- **η histograms (figures in appendix C):** in terms of the p_t difference, again bigger radii seem to improve the matching between the two levels. The peculiar behaviour of all algorithms in the range of $\eta = 1.5$ is due to the transition between the central and forward tracking detector. k_t, anti-k_t and Cambridge/Aachen exhibit a slightly better behaviour than the SISCone algorithm. In terms of the η difference, SIScone seems to have the best characteristics, as it shows the smallest deviation from 0 (especially for high η) and smaller dependencies on changes in R_0. In contrast, k_t, anti-k_t and Cam/Ac show higher dependencies on changes in R_0 and bigger deviations from 0, especially for high η. All algorithms exhibit a bigger η on had-level than on rec-level, but here only in the range of $\eta > 1$. This may also be related to the transition between central and forward tracking detector, which was already mentioned. This could explain the offset in η in the p_t histograms. The offset in ϕ is very similar for all algorithms, but not constant like in the p_t histograms. Instead, it is growing bigger from

$\eta \approx 1$ on.

- **ϕ histograms (figures in appendix D):** as for the p_t difference, a radius of $R_0 = 1.0$ seems to be the best choice for all algorithms. The k_t shows here and for the η difference the best performance. For the η differences, small radii improve the performance of all algorithms. The offset in ϕ is again very similar for all algorithms and radii.
- **M_{12}^2 histograms (figures in appendix E):** the behaviour seems to be very similar to the p_t plots. For M_{12}^2 less than 3000 GeV^2 the difference between hadron level and reconstructed level p_t is less than 0. In this region, smaller R_0 should be preferred for all jetfinders. In the region M_{12}^2 >3000 GeV^2 bigger radii yield better results. In terms of the η difference, SISCone again seems to be slightly less dependent on changes in R_0.

7.3 Comparison with the results of the DIS analysis

Comparing this work with B. Lemmers analysis on DIS data (4), one first recognizes the lower matching efficiencies in DIS. This is due to the lower p_t cuts (first jet p_t >7 GeV) compared to the cuts in photoproduction. Since lower p_t jets are less precisely defined, the efficiency decreases. The behaviour of the matching efficiency versus p_t and η seems to be similar. The ϕ plots differ because in the DIS analysis the jets are boosted into the Breit frame. The offset in ϕ in the difference plots is also present in the DIS work, which means that also there the ϕ on rec-level is bigger than the ϕ on had-level. Interestingly, the offset is smaller for Rapgap data. The other results are hard to compare because of the different cuts on p_t and different Q^2 range. In general, also for the jets in DIS k_t, anti-k_t and Cambridge/Aachen seem to be slightly preferred over the SISCone.

8. FINAL CONCLUSION AND OUTLOOK

The analysis shows that especially for k_t and anti-k_t $R_0 = 1.0$ is in most cases a good choice. SISCone in contrast works better with bigger radii. The behaviour of anti-k_t in terms of the jet matching is noticeable, here it shows the best efficiency, k_t in contrast shows the worst. The behaviour of Cambridge/Aachen is in most cases very similar to that of k_t and anti-k_t. One can conclude from that, that k_t, anti-k_t and Cambridge/Aachen are the preferred algorithms, all in all the differences between the jetfinders are rather small.

The mysterious offset in all ϕ difference histograms needs further investigations, a correction in the detector simulation and reconstruction process might be necessary.

Further studies of the differences between hadron and parton level and the separate consideration of direct and resolved photoproduction events might be interesting.

ACKNOWLEDGEMENTS

During my stay at DESY I had an interesting insight into scientific life and work. I want to thank Günter, Roman, Aziz, Adil and Zuzana from the H1 collaboration for all their help and the support during our work. Also thanks to my fellow summerstudent Boris Lemmer for the great entertainment during the work and life at DESY. Thanks to the fat, lazy cats lying all the time in front of the guesthouses and thus helping to create an enjoyable and inspiring atmosphere. Thanks also to the DESY snack and coffee machines for the reliable supply with all the things that are indispensable to life. Many thanks also to the organizers of the summerprogram, especially to Joachim Meyer and Andrea Schrader. Thanks to the lecturers for filling up our brains with huge amounts of physical knowledge. Finally I want to thank all the other summerstudents for the funny and pleasant time at DESY with all the parties, barbecues and city-tours. It was awesome! I like it!

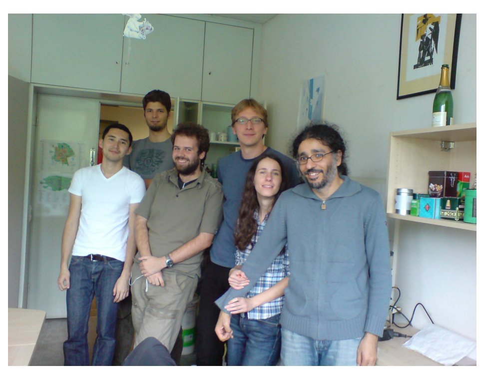

FIG 5. *My working group was really great!*

REFERENCES

[1] Matteo Cacciari and Gavin P. Salam. *The anti-k_t jet clustering algorithm.* JHEP 0804 (2008) 063 [arXiv:0802.1189]

[2] Gavin P. Salam and Gregory Soyez. *A practical seedless infrared-safe cone jet algorithm.* JHEP 05 (2007) 086 [arXiv:0704.0292]

[3] S. Caron. *Jets in photoproduction at HERA.* PhD thesis, RWTH Aachen 2002, DESY-THESIS-2002-035

[4] I. Strauch. *Jets with high transverse momenta in photoproduction at HERA.* PhD thesis, Universität Hamburg 2004, DESY-THESIS-2004-047

[5] B. Lemmer. *Comparing different jet-algorithms for DIS at high Q^2.* DESY summer school 2008 student report.

[6] Torbjorn Sjostrand et. al. *PYTHIA 6.4 physics and manual.* JHEP 05 (2006) 026

APPENDIX A: JET MATCHING PLOTS

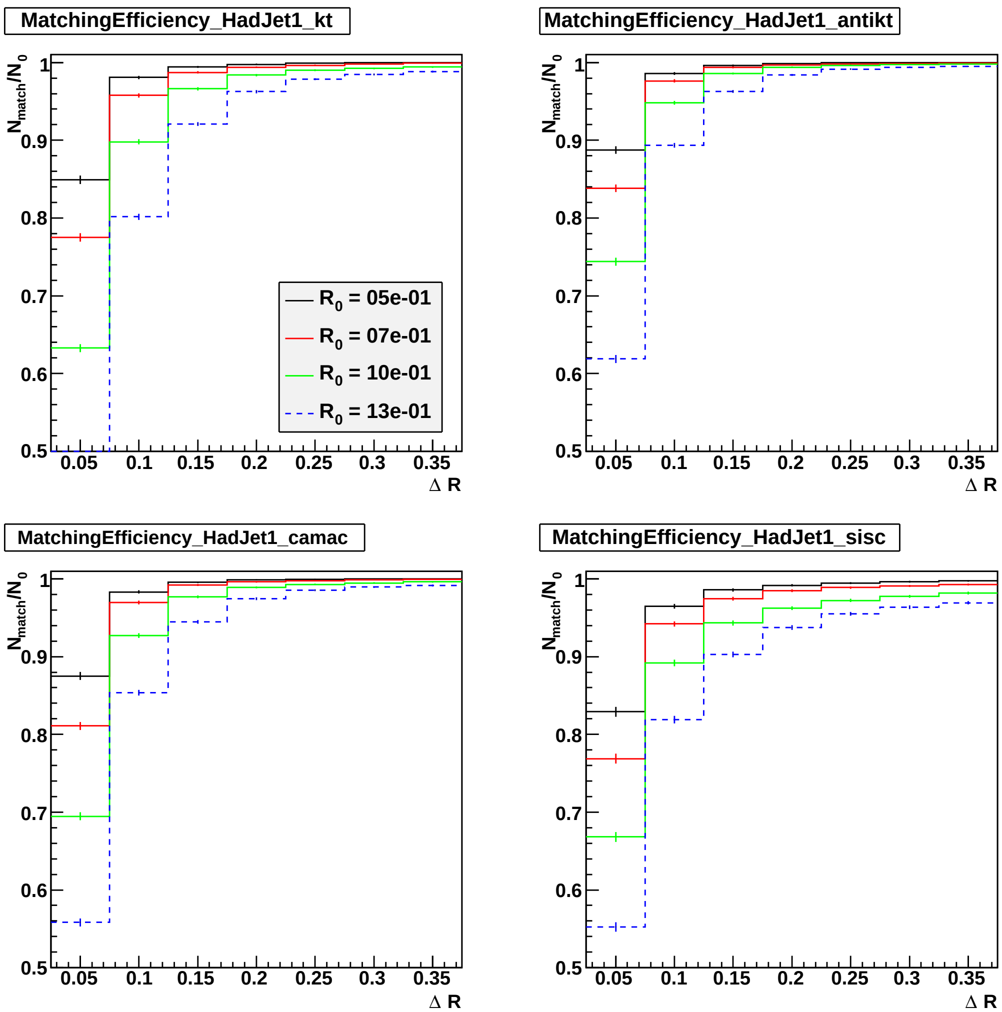

FIG 6. *1st jet on hadron level: matching efficiency for all jet-algorithms with different radius parameters (colored lines). The k_t algorithm seems to be more sensitive to changes in R_0 than the anti-k_t algorithm. The errors are calculated according to a binomial distribution of N_{match}.*

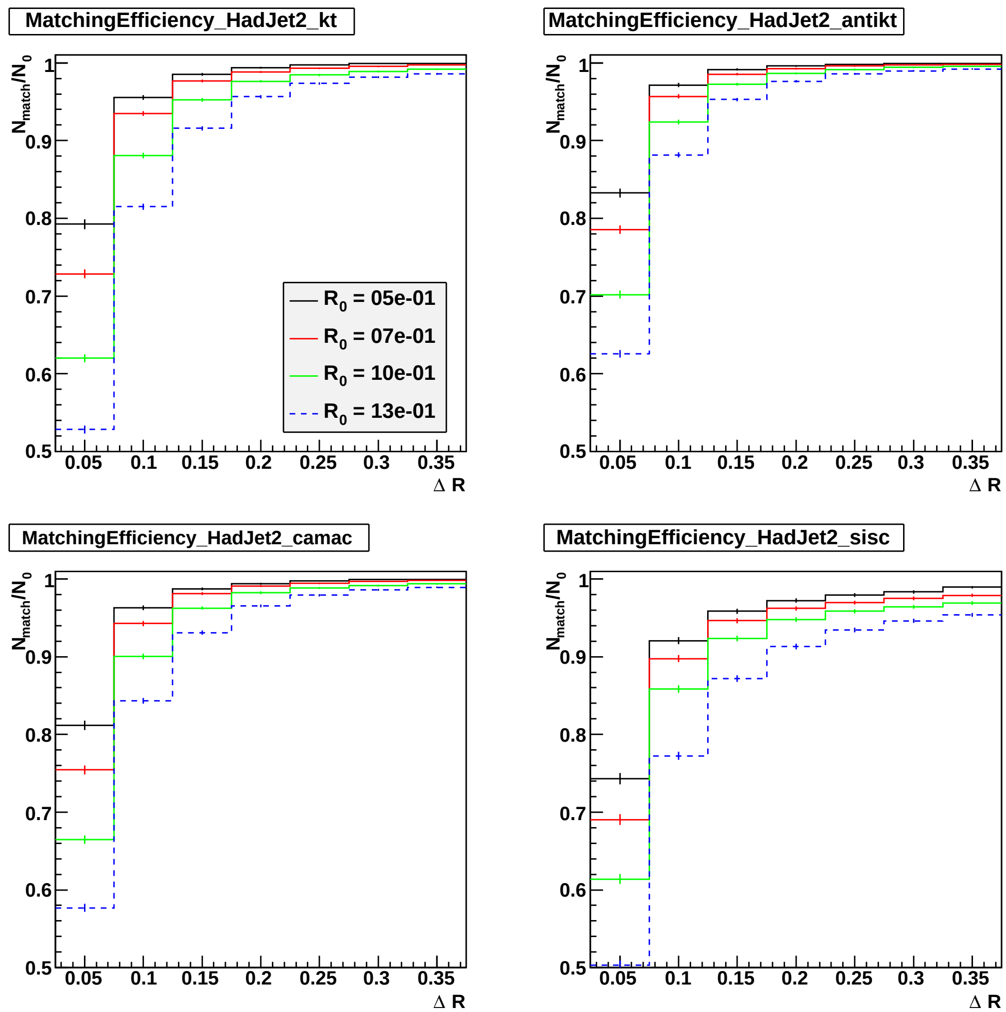

FIG 7. *2nd jet on hadron level: matching efficiency for all jet-algorithms with different radius parameters (colored lines). The errors are calculated according to a binomial distribution of* N_{match}.

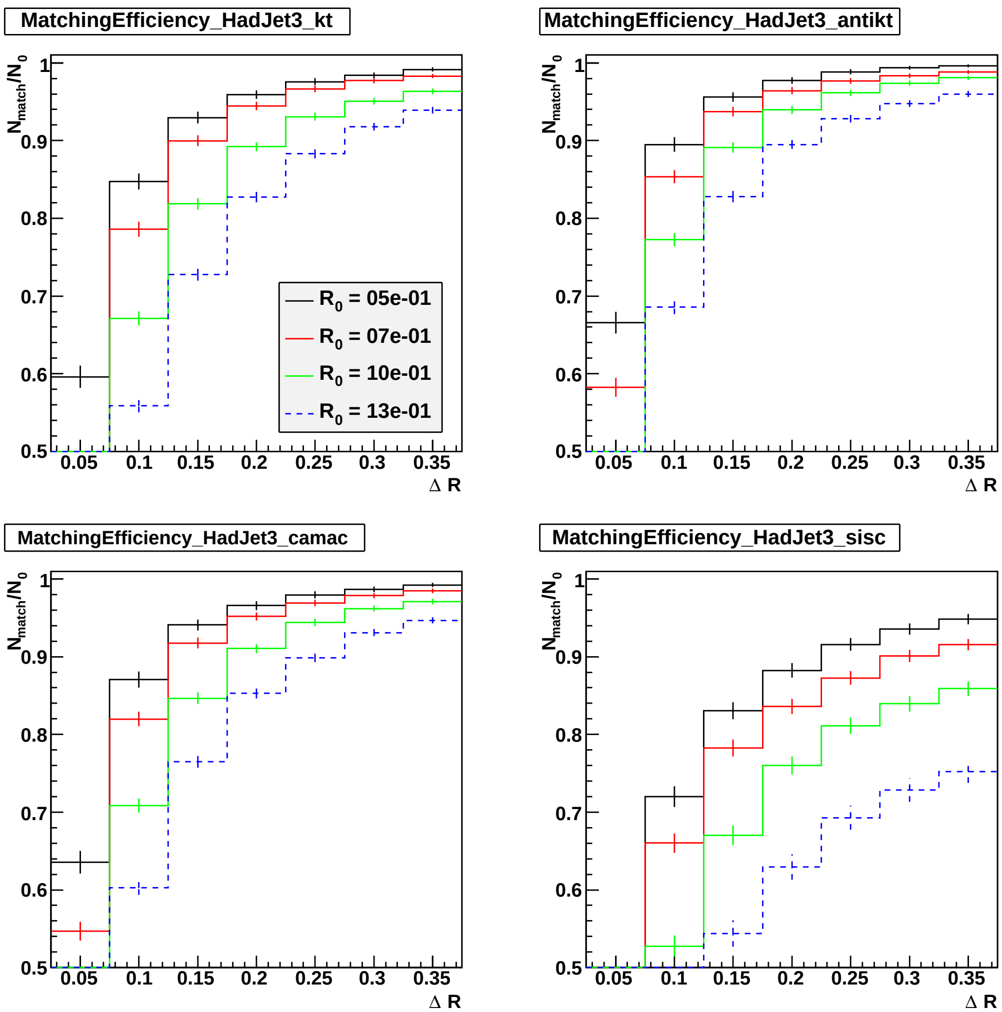

FIG 8. *3rd jet on hadron level: matching efficiency for all jet-algorithms with different radius parameters (colored lines). The errors are calculated according to a binomial distribution of* N_{match}.

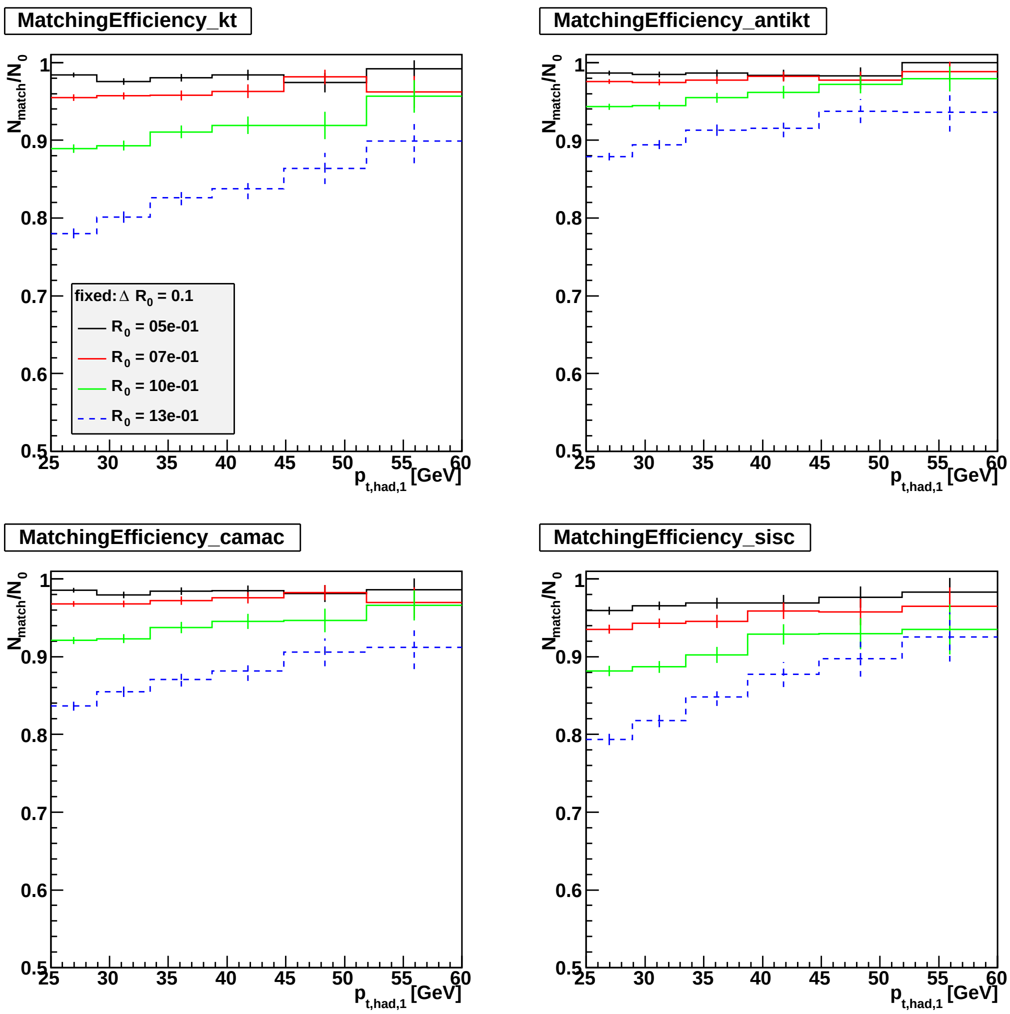

FIG 9. *1st jet on hadron level: matching efficiency for all jet-algorithms with different radius parameters (colored lines). The parameter $\Delta R = 0.1$ is here fixed, the matching efficiency is plotted versus p_t of the 1st had-jet. The errors are calculated according to a binomial distribution of N_{match}.*

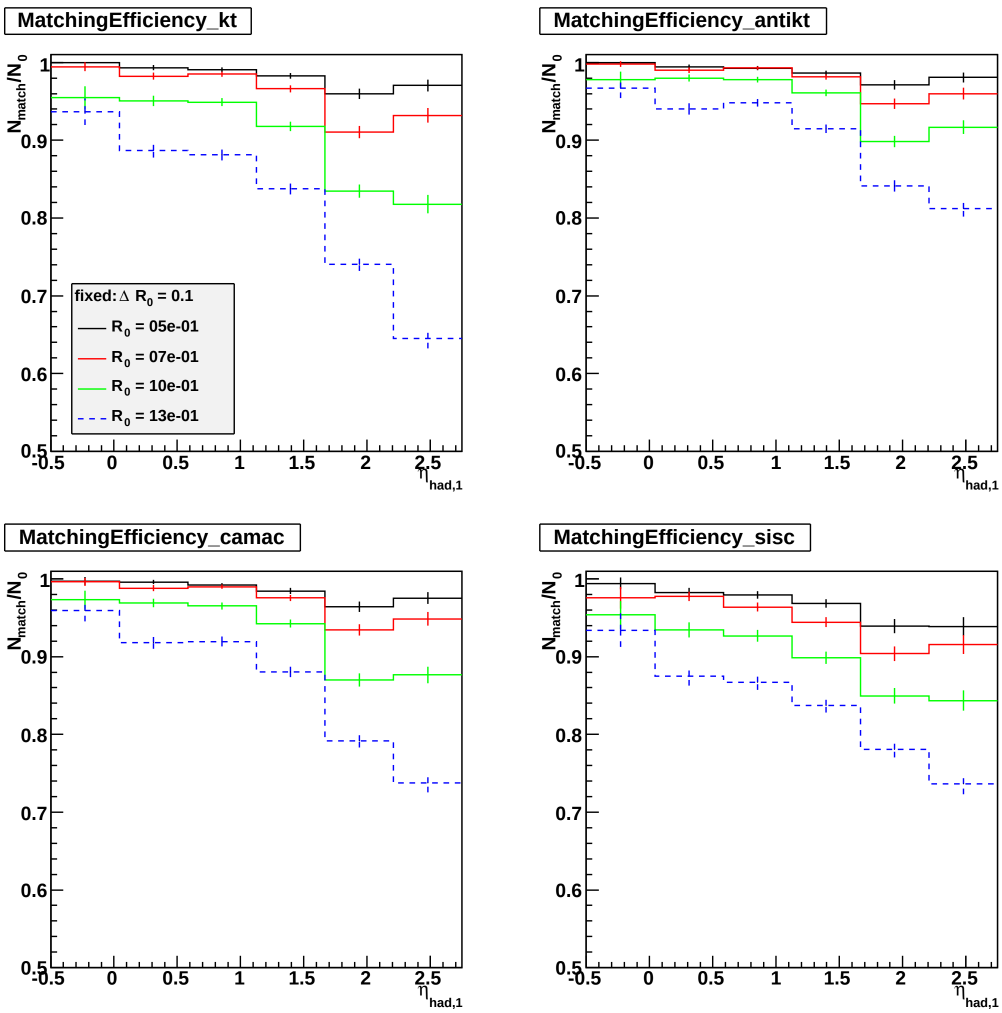

FIG 10. *1st jet on hadron level: matching efficiency for all jet-algorithms with different radius parameters (colored lines). The parameter* $\Delta R = 0.1$ *is here fixed, the matching efficiency is plotted versus* η *of the 1st had-jet. The errors are calculated according to a binomial distribution of* N_{match}.

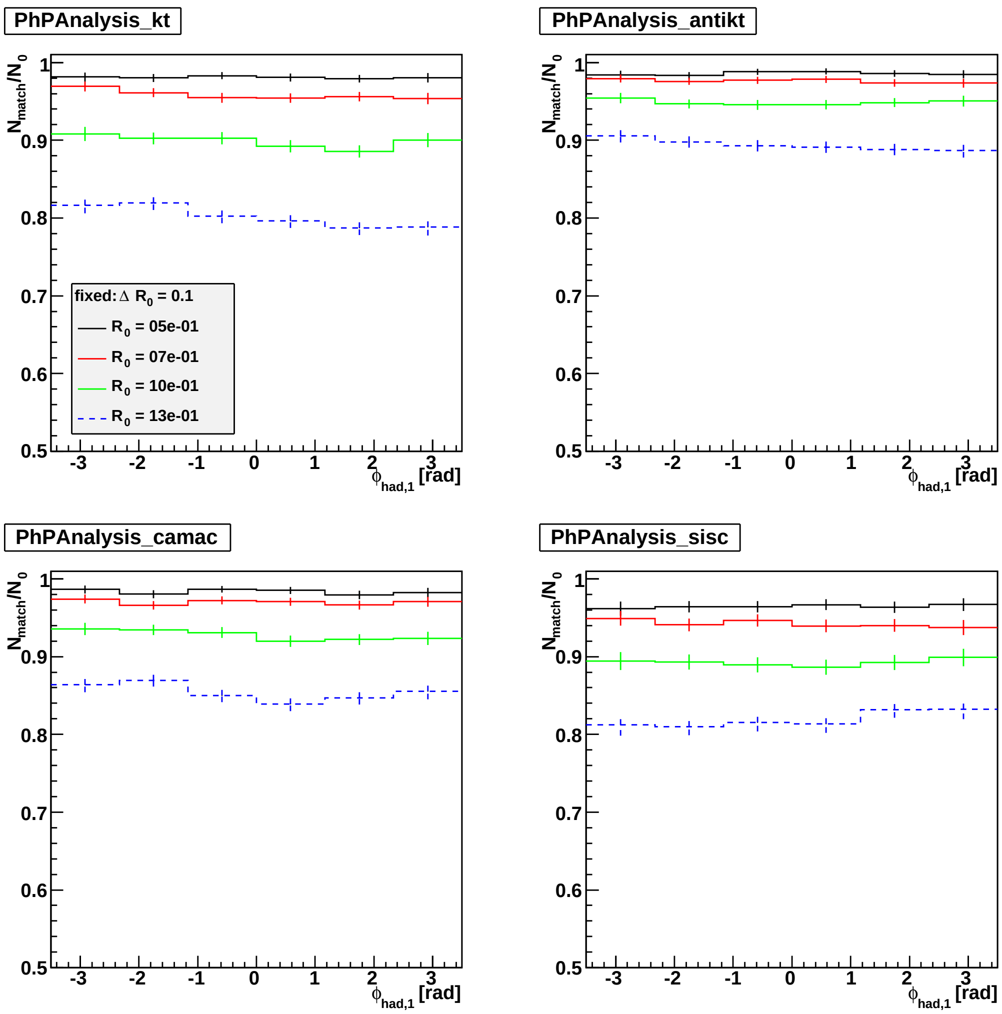

FIG 11. *1st jet on hadron level: matching efficiency for all jet-algorithms with different radius parameters (colored lines). The parameter* $\Delta R = 0.1$ *is here fixed, the matching efficiency is plotted versus* ϕ *of the 1st had-jet. The errors are calculated according to a binomial distribution of* N_{match}.

APPENDIX B: REC-HAD-LEVEL DIFFERENCES VERSUS $P_{T,HAD,1}$

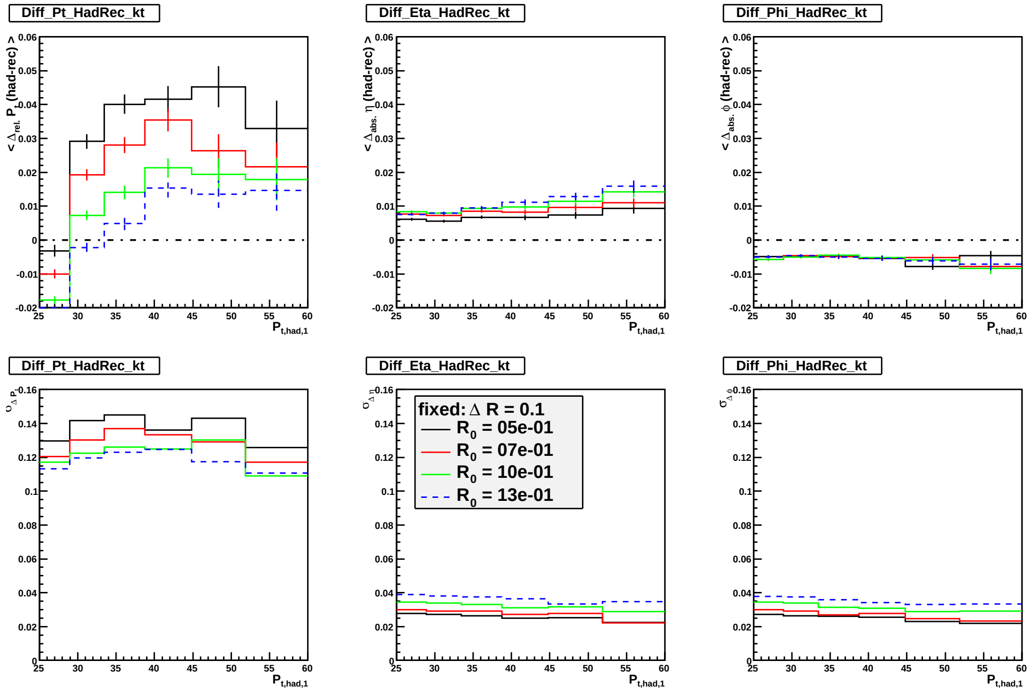

FIG 12. *In the first row, the mean values of the difference distributions between had- and rec-level are shown versus p_t of the first jet at had-level. The second row shows the associated sigmas of the difference distributions. The first column shows the difference in p_t, the second one the difference in η and the third one the difference in ϕ. Note that for the two angular quantities the absolute differences between had- and rec-level were calculated, whereas the p_t difference is calculated relative to the had-level p_t of the jet. Used jet-algorithm is k_t.*

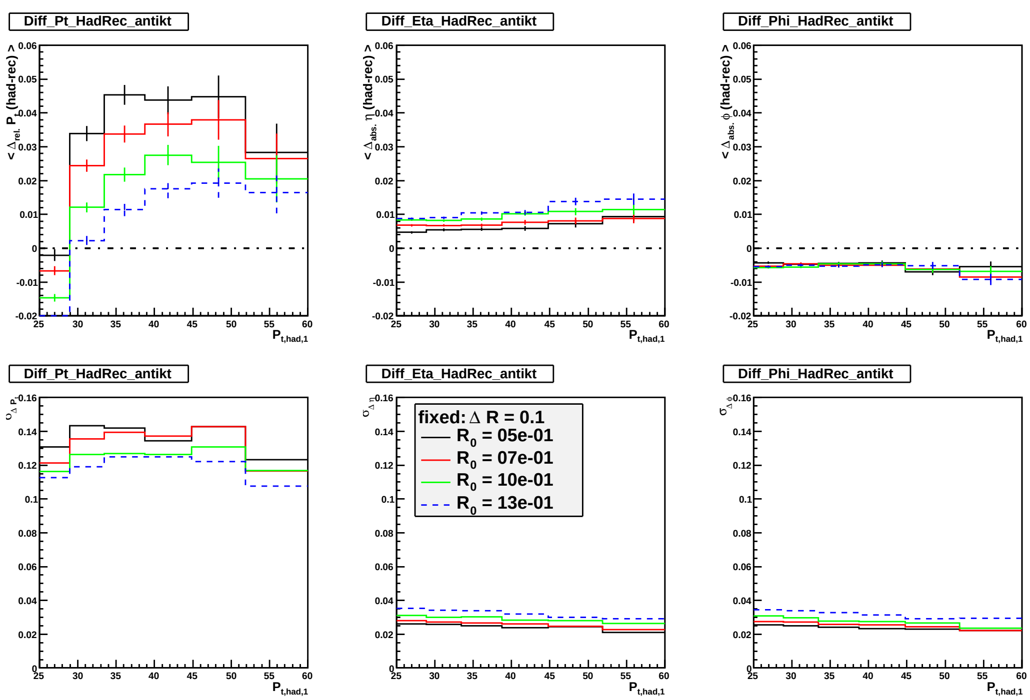

FIG 13. *Used jet-algorithm is anti-k_t.*

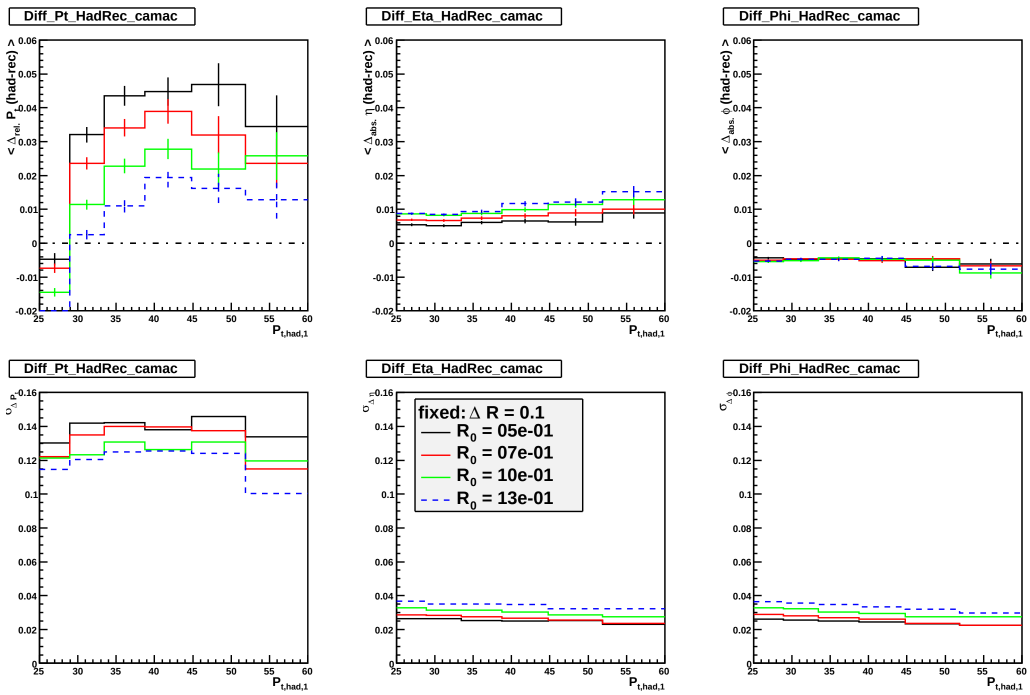

FIG 14. *Used jet-algorithm is Cambridge/Aachen.*

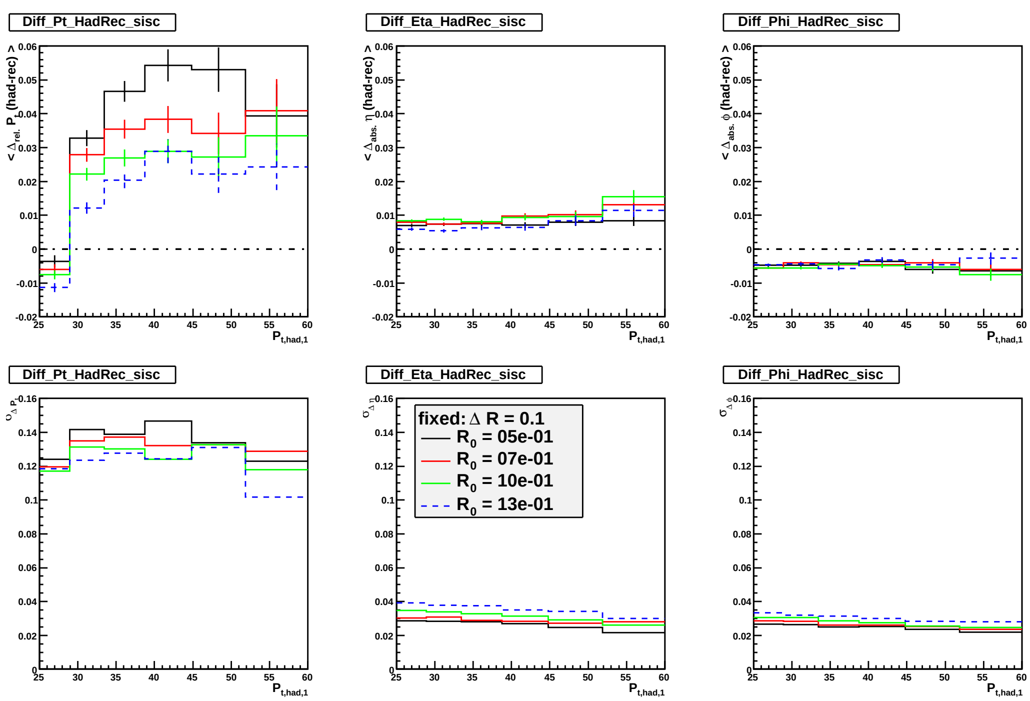

FIG 15. *Used jet-algorithm is SISCone.*

APPENDIX C: REC-HAD-LEVEL DIFFERENCES VERSUS $\eta_{HAD,1}$

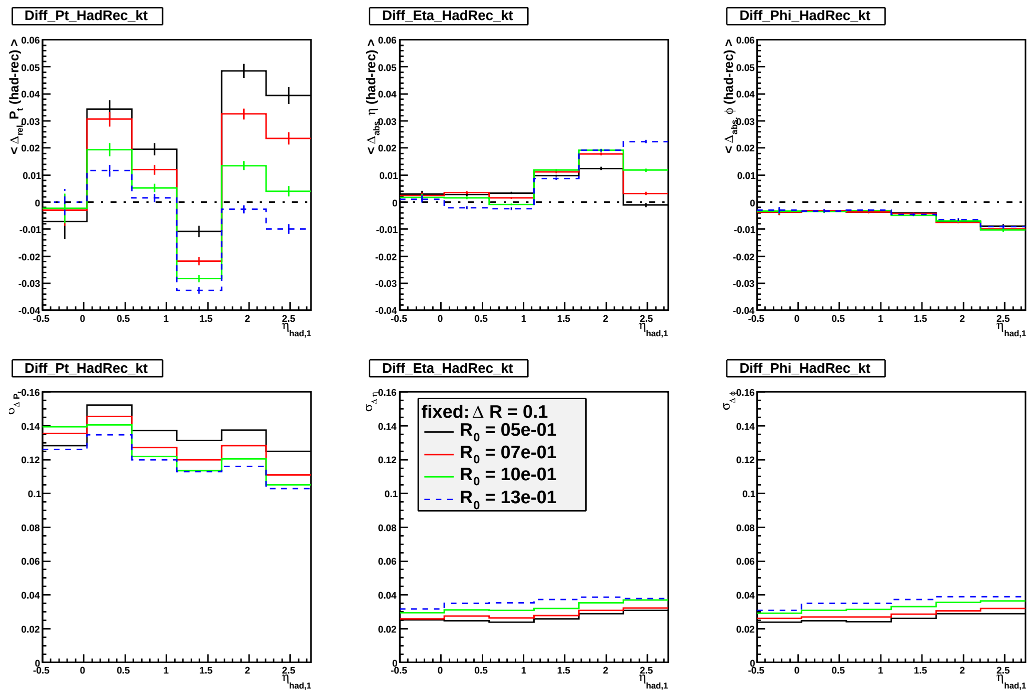

FIG 16. *In the first row, the mean values of the difference distributions between had- and rec-level are shown versus η of the first jet at had-level. The second row shows the associated sigmas of the difference distributions. The first column shows the difference in p_t, the second one the difference in η and the third one the difference in ϕ. Note that for the two angular quantities the absolute differences between had- and rec-level were calculated, whereas the p_t difference is calculated relative to the had-level p_t of the jet. Used jet-algorithm is k_t.*

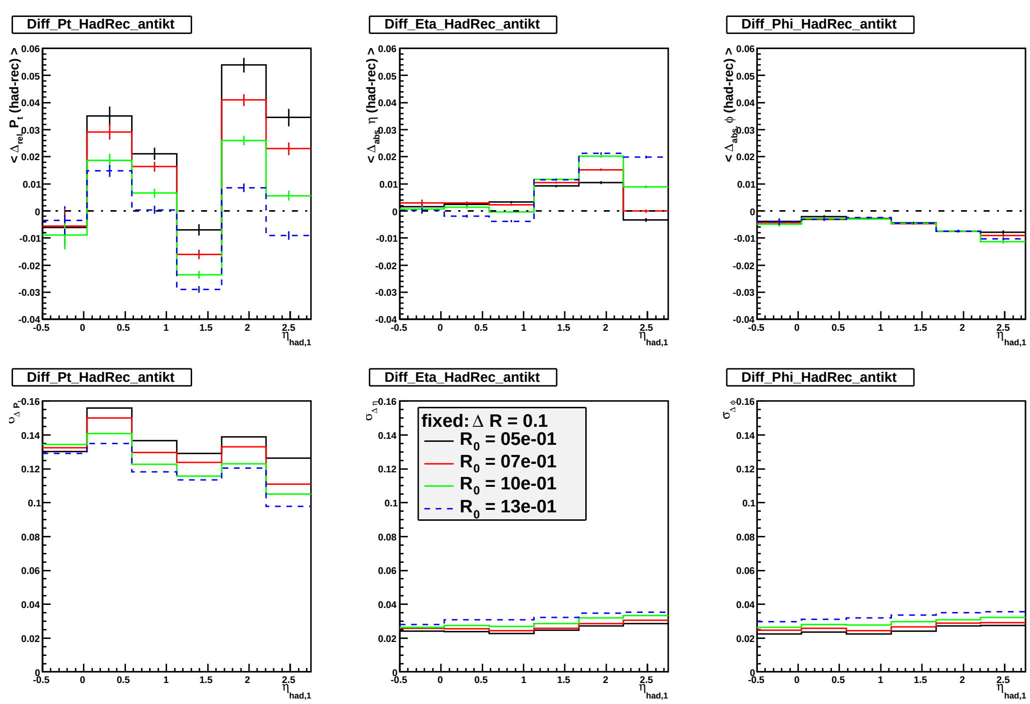

FIG 17. *Used jet-algorithm is anti-k_t.*

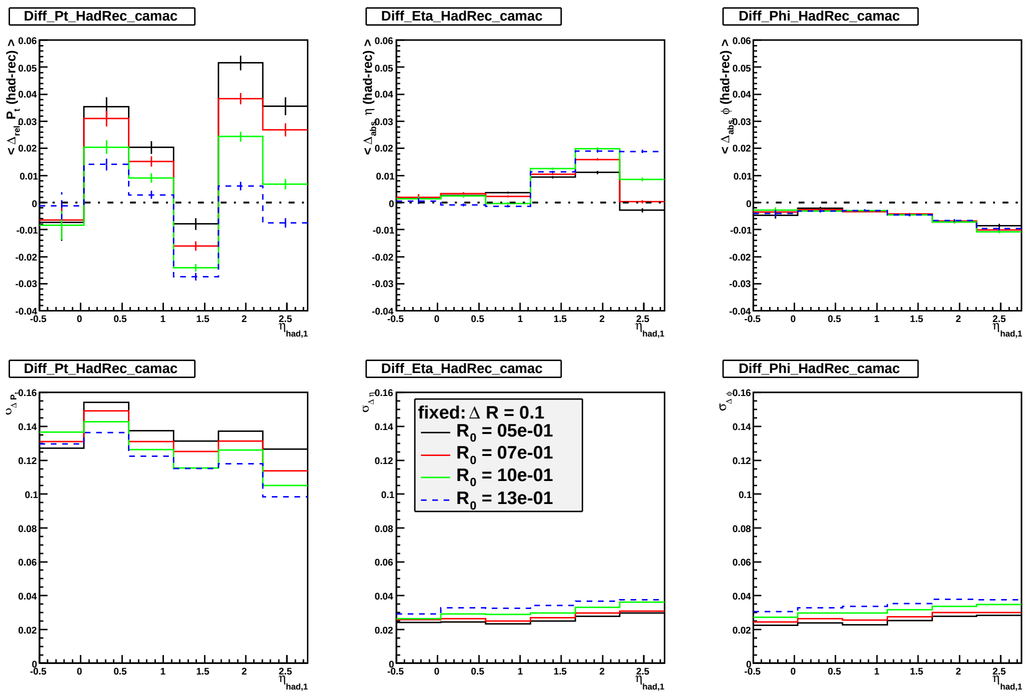

FIG 18. *Used jet-algorithm is Cambridge/Aachen.*

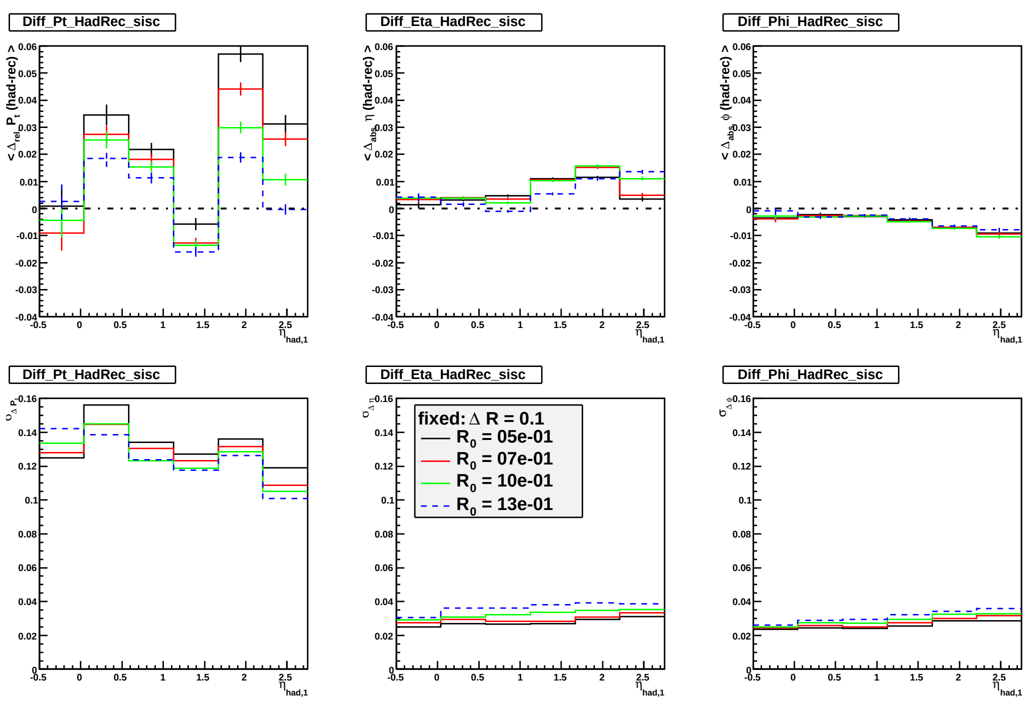

FIG 19. *Used jet-algorithm is SISCone.*

APPENDIX D: REC-HAD-LEVEL DIFFERENCES VERSUS $\phi_{HAD,1}$

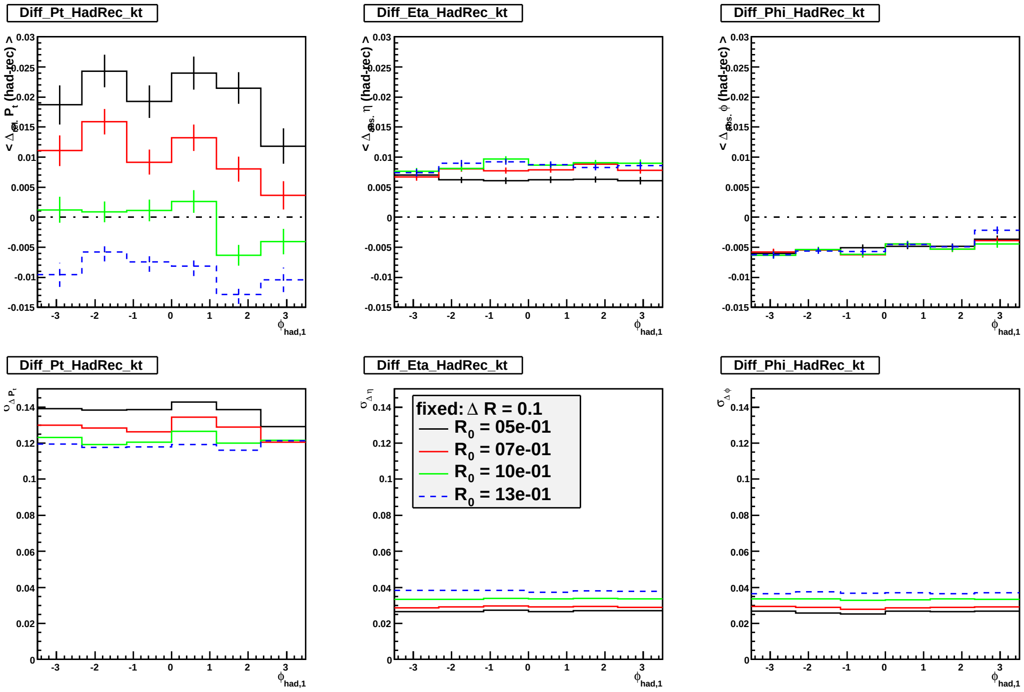

FIG 20. *In the first row, the mean values of the difference distributions between had- and rec-level are shown versus ϕ of the first jet at had-level. The second row shows the associated sigmas of the difference distributions. The first column shows the difference in p_t, the second one the difference in η and the third one the difference in ϕ. Note that for the two angular quantities the absolute differences between had- and rec-level were calculated, whereas the p_t difference is calculated relative to the had-level p_t of the jet. Used jet-algorithm is k_t.*

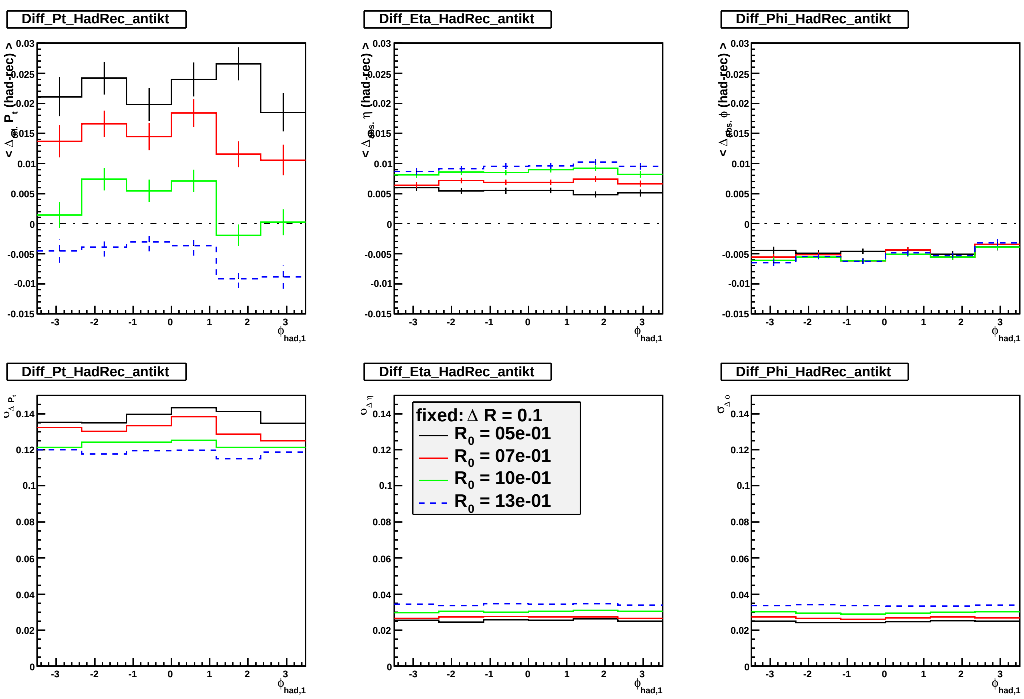

FIG 21. *Used jet-algorithm is anti-k_t.*

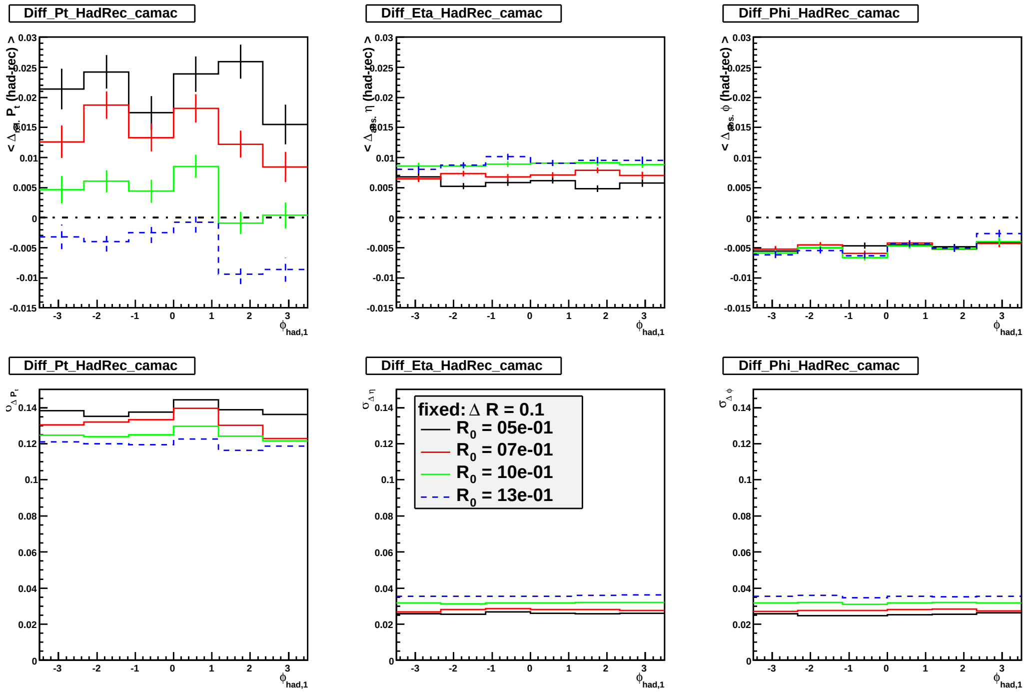

FIG 22. *Used jet-algorithm is Cambridge/Aachen.*

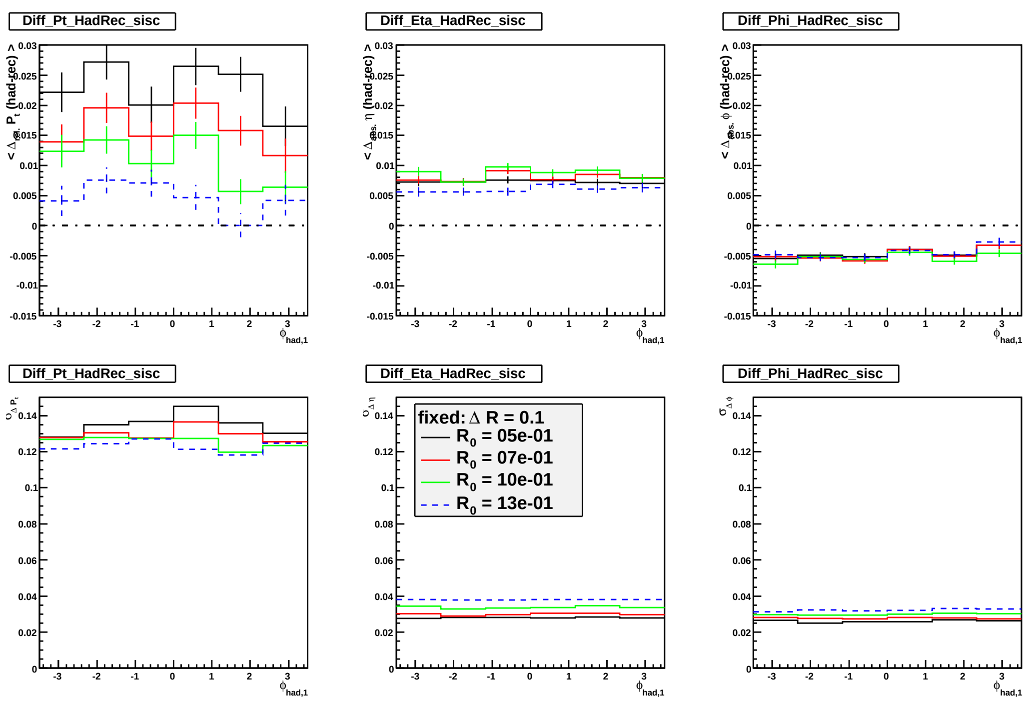

FIG 23. *Used jet-algorithm is SISCone.*

APPENDIX E: REC-HAD-LEVEL DIFFERENCES VERSUS $M^2_{HAD,12}$

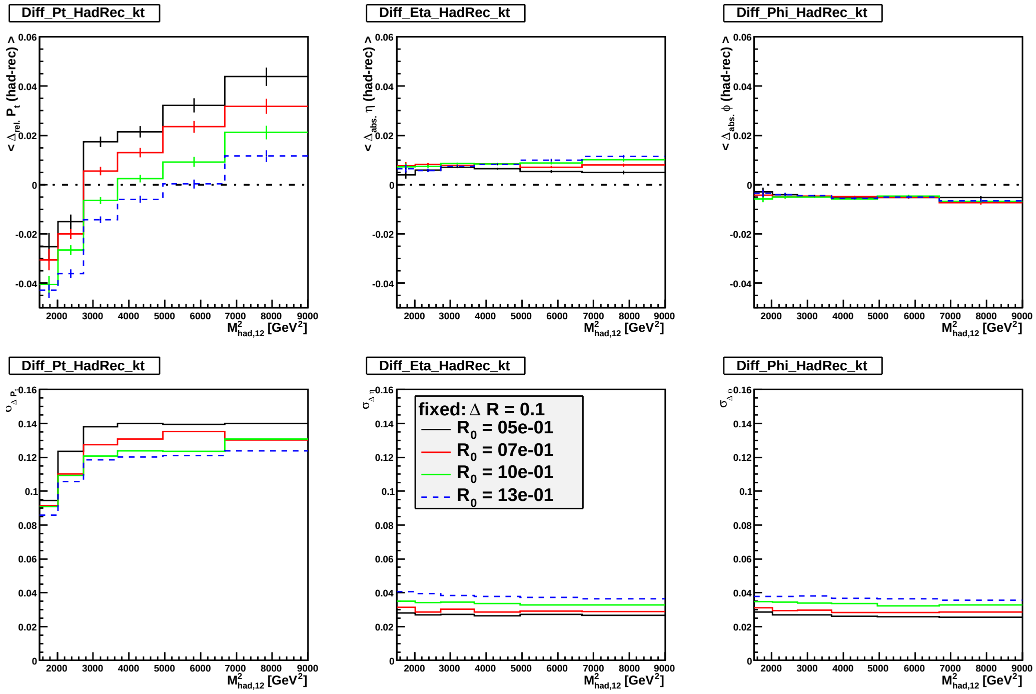

FIG 24. *In the first row, the mean values of the difference distributions between had- and rec-level are shown versus M_{12}^2 of the first jet at had-level. The second row shows the associated sigmas of the difference distributions. The first column shows the difference in p_t, the second one the difference in η and the third one the difference in ϕ. Note that for the two angular quantities the absolute differences between had- and rec-level were calculated, whereas the p_t difference is calculated relative to the had-level p_t of the jet. Used jet-algorithm is k_t.*

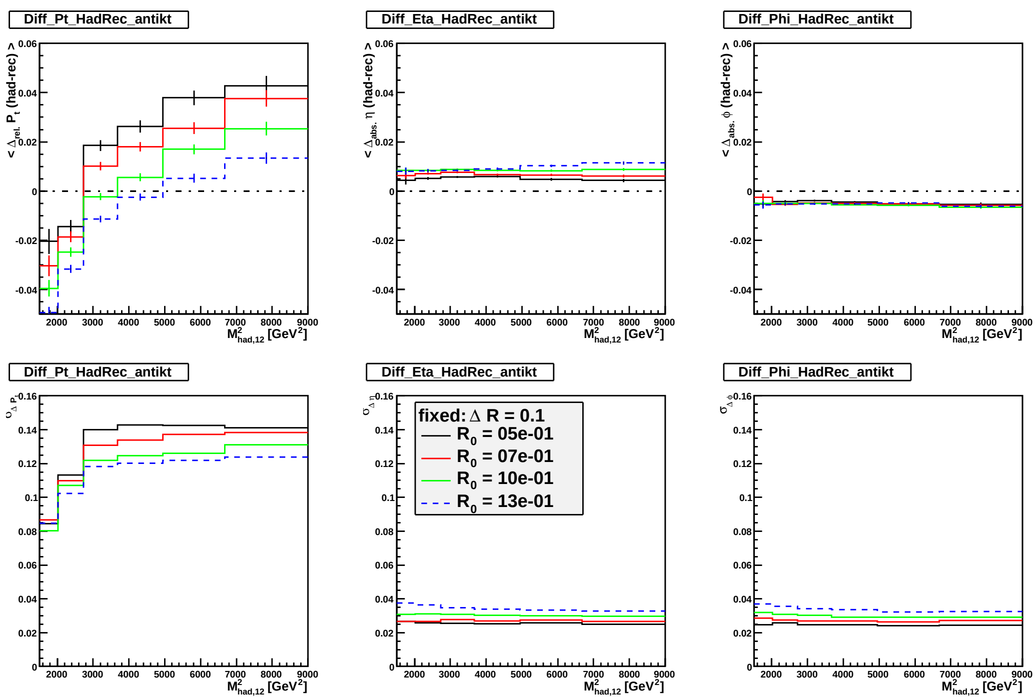

FIG 25. *Used jet-algorithm is anti-k_t.*

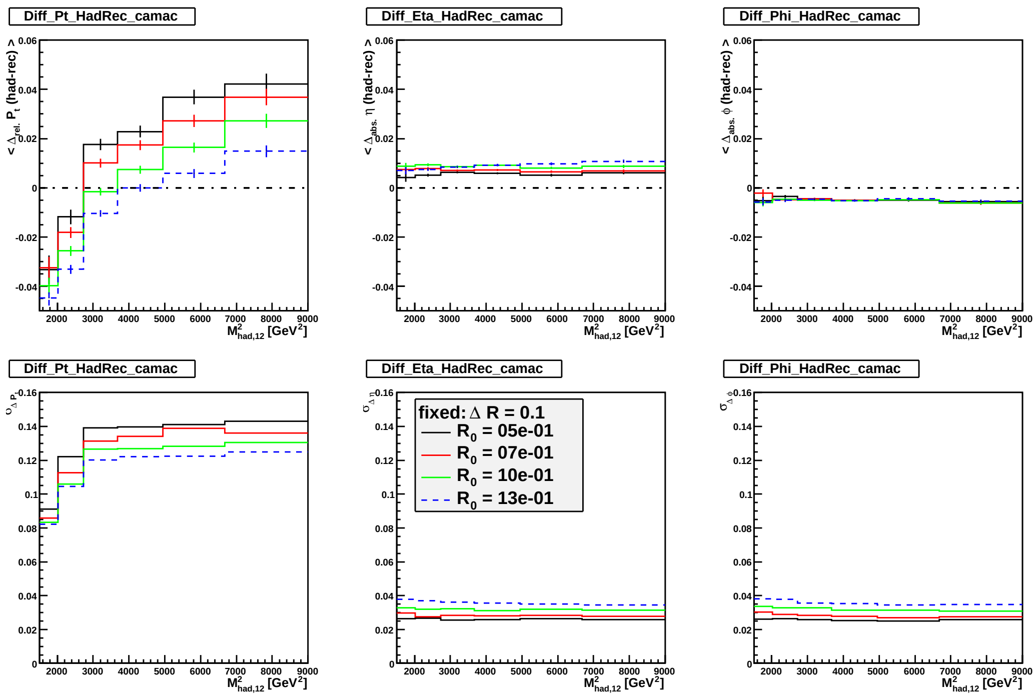

Fig 26. *Used jet-algorithm is Cambridge/Aachen.*

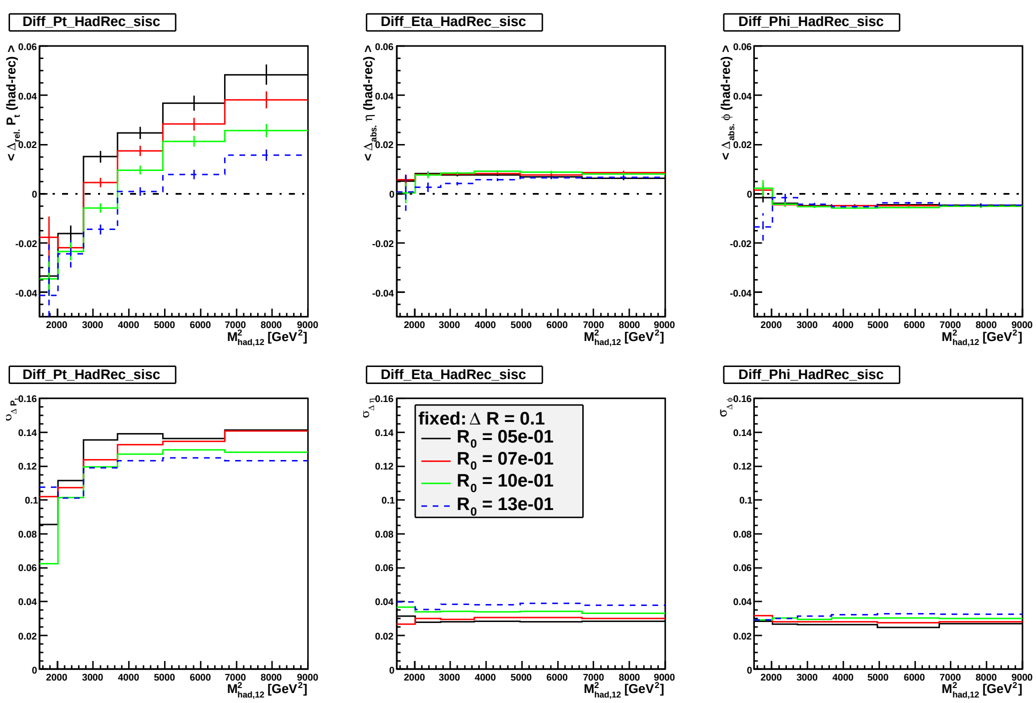

Fig 27. *Used jet-algorithm is SISCone.*