30 August 2024
esterete berga@gmail.com

Hard radiation in top quark decays

Estere Agnese Teberga
CERN, CH-1211 Geneva, Switzerland

Supervisors: Markus Seidel, Martijn Mulders, Karlis Dreimanis

Keywords: CERN Summer Student project report

Summary

Due to concerns regarding the gluon emission recoil in top quark modelling, it's been suggested to treat the top quark as recoiler in $t \to bW$ decay. To test this as a possibility to reduce the top mass measurement uncertainty, an analysis is done on the $t\bar{t}$ semileptonic decay channel by looking for hard radiation outside b-jet cones and reconstructing it as either an additional soft jet (for hadronic decay) or as a fat b-jet (for leptonic decay) with the goal to find differences between RTT and mass or FSR variations. The soft jet study shows no significant differences for RTT, but succeeded in enriching the sample in FSR events. The fat jet substructure study shows promise for the subjet ΔR observable, showing distinct shape for the RTT sample. A fast simulation of CMS detector effects was applied, demonstrating that this should also be measurable in CMS data.

1 Introduction and background

The top quark is the heaviest Standard Model particle, and a precise measurement of its mass has implications for multiple open questions in the high energy particle physics community. The most recent m_t measurement has an uncertainty of 0.37 GeV [1], and a contribution to this uncertainty is partially due to radiation outside jet cones. The top quark decays predominantly to a W boson and b quark. The W boson can decay hadronically (into 2 quarks) or leptonically (into a charged lepton and its corresponding neutrino), and the quarks in these decays turn into jets of particles. Current modelling of these processes have raised concerns over unconstrained energy loss outside of the jet cones.

One proposed remedy for this problem is adjusting the gluon emission recoil scheme from top quark decay. Recoil in the context of high energy physics refers to the redistribution of energy and momentum in a decay, and it can have significant effects on the reconstruction of invariant mass. While previous releases of Pythia8 have designated either the W boson or the b quark as the recoiler in top quark decays, this modelling has raised concerns of having an unphysical representation of out-of-cone radiation and phase space. This may be solved by designating the top quark as the recoiler (RTT) [2, 3]. A previous study, during which $t\bar{t}$ Monte Carlo samples with RTT were made and the top quark mass was extracted, showed that the effects of RTT had a non-negligible impact on the top mass in a lepton+jets analysis, causing a mass shift of at least 200 MeV, which is on the same scale as experimental uncertainties in Run 2 top mass measurements [4]. Due to the reduced statistical uncertainty for this mass measurement, this impact is significant enough to call for further inquiry into RTT effects.

The goal of this study is to constrain the additional radiation in top quark decay by identifying an additional jet from hard out-of-cone radiation. This is done by analyzing simulated semileptonic $t\bar{t}$ events and looking for hard radiation from the b-jet in both decay branches (Fig. 1). To test the performance of this treatment, several observables are measured with the aim to check for differences between the Monte Carlo simulated RTT sample and the default sample, as well as to see if RTT is distinguishable from top mass variations and final state radiation (FSR) variations.

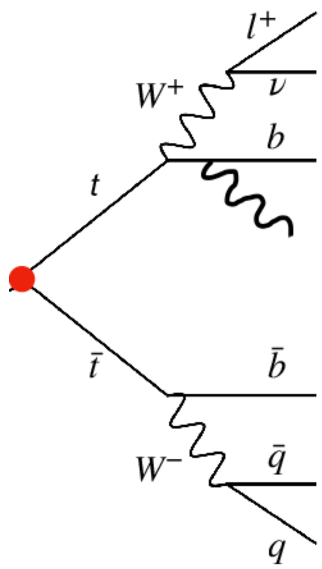

Figure 1: Feynman diagram of $t\bar{t}$ semileptonic decay with additional jet

2 Event selection

The analysis is implemented in a Rivet [5] routine called MC_TOP_FSR, found in merge request 260, and it consists of two independent parts - one is analyzing the hadronic decay chain, while the other is focused at the leptonic decay part. To select events consistent with semileptonic decay channel expectations, only events with exactly one high p_T lepton are selected, as well as at least 4 jets - 2 b-tagged and 2 untagged. A general cut on jet parameters is imposed for both analyses:

- Jet cut: $\eta < 2.4$, $p_T > 10$ GeV

Then the jets are sorted into categories of b-jet, light jet and soft jet as follows:

- If jet $p_T > 30$ GeV and is b-tagged $\rightarrow$ b-jet,
- If jet $p_T > 30$ GeV and is not b-tagged $\rightarrow$ light jet (l-jet),
- If $10 < p_T < 30$ GeV $\rightarrow$ soft jet.

To reconstruct the top quarks, a minimising function is used. This function takes a parameter K, which is the quadratic sum of W-boson candidate mass difference to true W mass and the difference between leptonic top and hadronic top candidate masses. K is minimized by iterating through all b-jets and l-jets and constructing hadronic W (from 2 l-jets), hadronic t (from hadronic W and a b-jet) and leptonic t (from leptonic W and a b-jet. When K is minimum, the hadronic top and leptonic top candidates are declared as the reconstructed $t\bar{t}$.

For the hadronic decay side a soft jet study is done. The soft jet analysis histograms are filled if there is at least one soft jet and its angular distance from the b-jet is ΔR < 1. From this a 4-jet top quark invariant mass is reconstructed.

For the leptonic decay side study the events are clustered with fat jets with $R = 0.8$ and then the fat jet that is closest to the b-jet is analyzed. On the charged constituent particles of the fat jet a cut of $p_T > 1$ GeV is applied.

The analysis code is used on 4 MC samples - the central sample with $m_t = 172.5$ GeV, mass variations of $m_t = 173.5$ GeV, $m_t = 171.5$ GeV, and the RTT sample. Additional samples are extracted from the central sample with final state radiation variations via event weights - FSR up and FSR down.

3 Hadronic decay results

The initial idea for this analysis was to check the two dimensional distribution of the mass of the top quark m_t with the mass of the reconstructed top quark from four jets m_{t+j}, including the soft jet from the b-quark. The expected result would be to see two separate populations in this distribution, in hope to differentiate between initial state radiation processes (distributed in a vertical line) and final state radiation (distributed in a horizontal line). However, looking at the plots (Fig. 2, 3) shows that this is not the case.

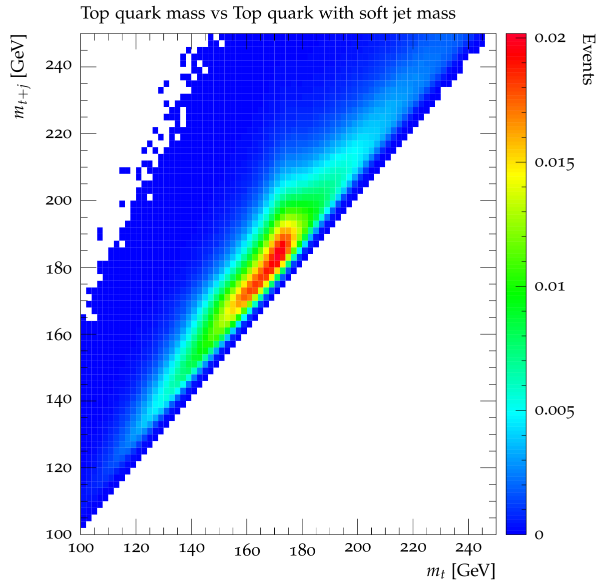

Figure 2: 2D histogram of central sample for m_t vs m_{t+j}

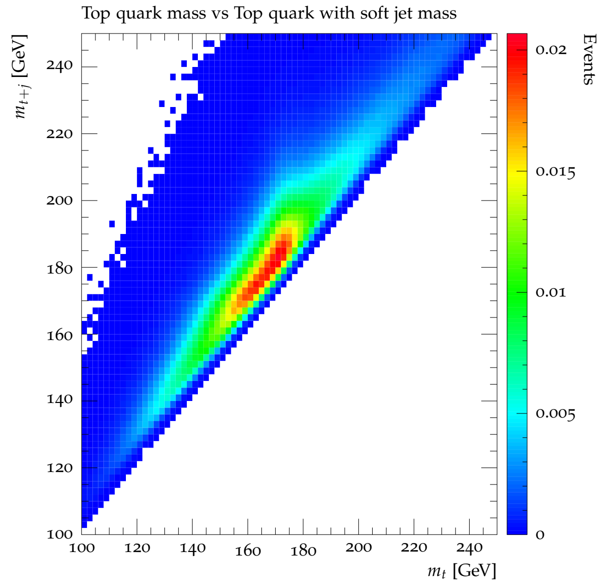

Figure 3: 2D histogram of RTT sample for m_t vs m_{t+j}

For closer inspection we look into several one dimensional distributions of different slices of the 2D histogram. The following plots were made from the full available datasets of the central, $m_t = 171.5$ GeV, $m_t = 173.5$ GeV and RTT Monte Carlo samples for maximum statistics. Fig. 4 was made by just filling in the m_{t+j} in a regular histogram, while figures 5 and 6 were made by filling histograms with m_{t+j} with selecting events of $m_{t+j} < 170$ GeV and $m_t + m_{t+j} < 345$ GeV, respectively. Initial state radiation (ISR) variations for these plots were also checked, but yielded a change of 2% at most.

The ratio plots for these distributions show that the RTT variation is not definitively distinguishable from from the top mass variations, concluding that this method does not help reduce uncertainty for top mass measurement.

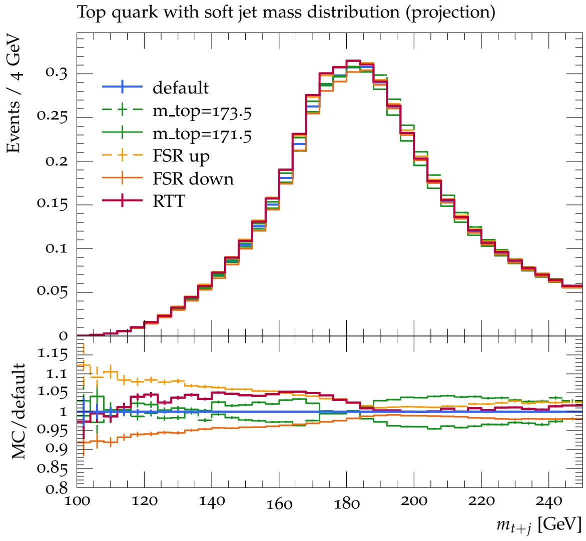

Figure 4: Projection of m_{t+j} on the y axis

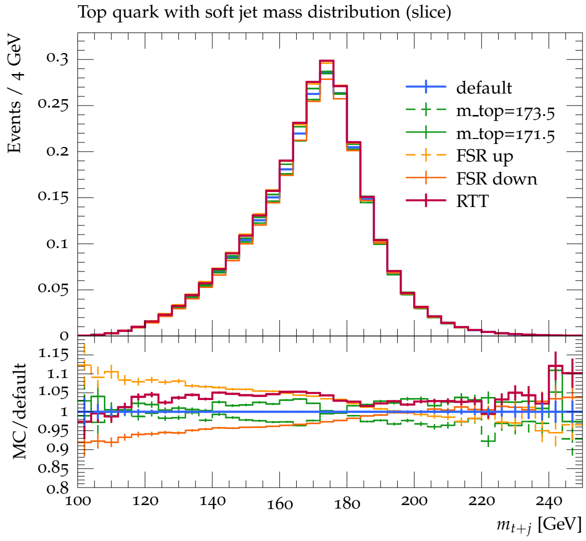

Figure 5: Slice of m_{t+j}

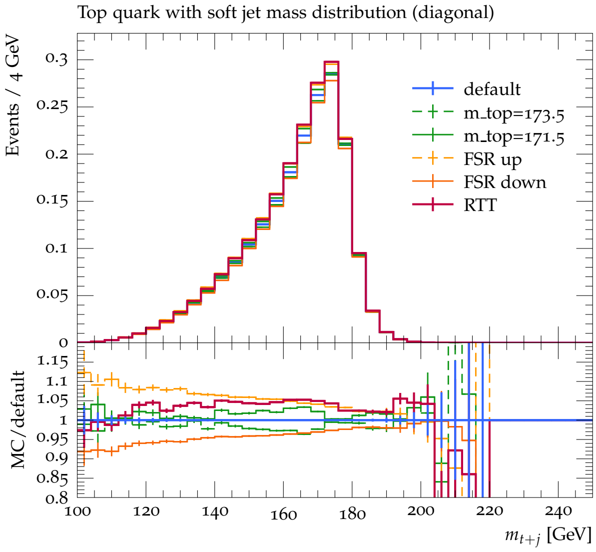

Figure 6: Diagonal slice of m_{t+j}

4 Leptonic decay results

For the analysis of the leptonic decay side, a fat jet of $R = 0.8$ around the b-jet was selected by requiring the jet to be b-tagged and having an angular distance of $\Delta R < 0.4$ to the original leptonic b-jet. This jet was then reclustered with the Cambridge-Aachen algorithm for better jet grooming and jet substructure studies. Jet substructure analysis plots were made from 1) just the electrically charged constituents and 2) all constituent particles. The chosen observables for this analysis are N-subjettiness ratios τ_{21}, τ_{32}, τ_{43} as well as momentum fraction z_g and subjet ΔR after soft-drop declustering.

In the N-subjettiness ratio plots the RTT sample seems to either follow a similar distribution to the top mass variations or the FSR down variation in the high-statistics regions, so no conclusions can be made just from these plots. An additional study for the feasibility of making NanoAOD files was done, by selecting only jets with $p_T > 170$ GeV. Reviewing these distributions showed that NanoAOD files are unfeasible for this analysis, since for NanoAOD we would need the prompt lepton from the W decays to be excluded from the jet clustering to see the QCD effects.

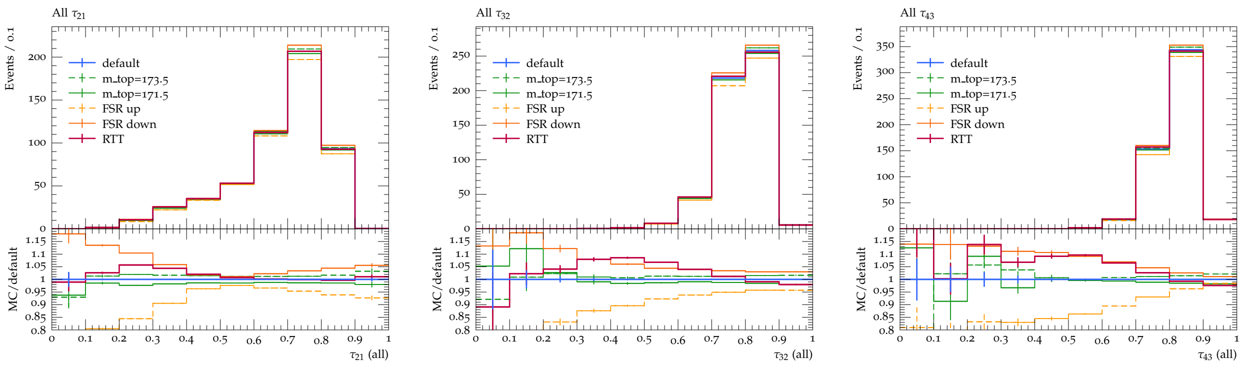

Figure 7: N-subjettiness ratios from all constituent particles

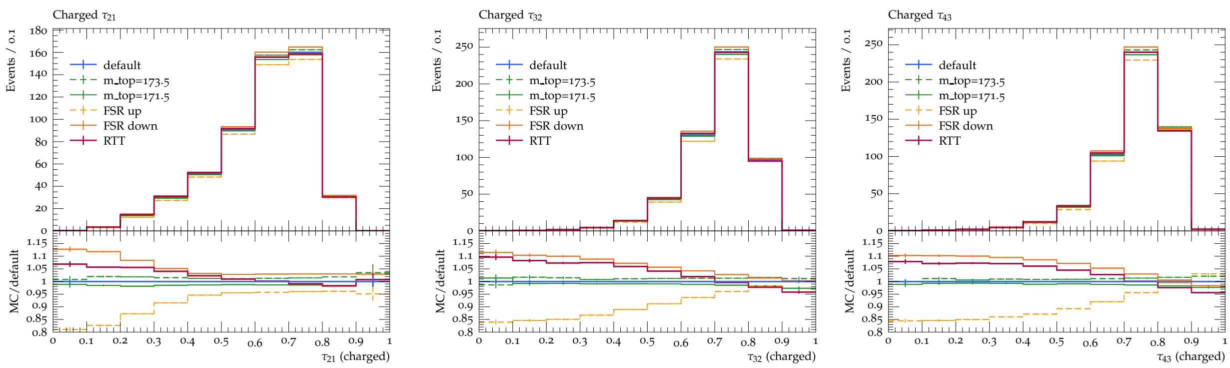

Figure 8: N-subjettiness ratios from charged constituent particles

For the momentum fraction z_g and subjet ΔR plots a jet soft-drop declustering algorithm was applied [6]. Soft drop declustering is a jet grooming procedure which recursively removes wide-angle soft radiation from a jet. For this analysis the soft drop threshold was chosen as $z_{cut} = 0.1$ and angular exponent $\beta = 0$, meaning that during declustering, when the subjet momentum is more than 10% of its parent jet, the algorithm stops and the values of z_g and subjet ΔR are returned.

The soft-drop observables are largely independent from top quark mass, so it is useful to look for differences to FSR variations. While the RTT sample for momentum fraction plots (Fig. 9, 11) seems to be very similar to the central sample, in the subjet ΔR plots it seems to differ from all other samples, for subjets from both charged and all constituents. Specifically, the narrow angle radiation $\Delta R < 0.4$ for RTT seems to follow a similar declining trend as FSR down in the ratio plot, but for wider angle radiation it gets larger than for all other samples, showing a distinct shape for RTT. This could hint at some new effects from RTT that differ from both FSR and mass variations, which means that this subjet ΔR observable could help distinguish between different models.

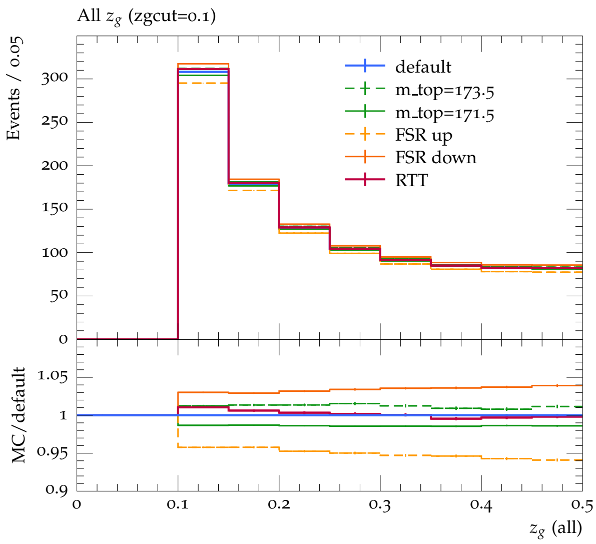

Figure 9: Momentum fraction z_g (all constituents)

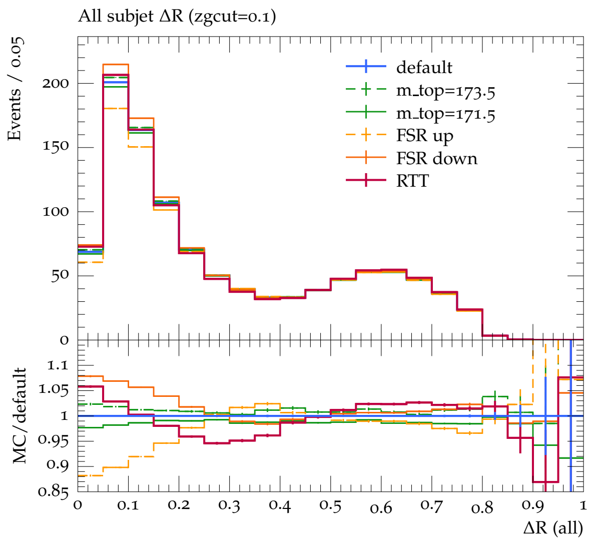

Figure 10: Subjet ΔR (all constituents)

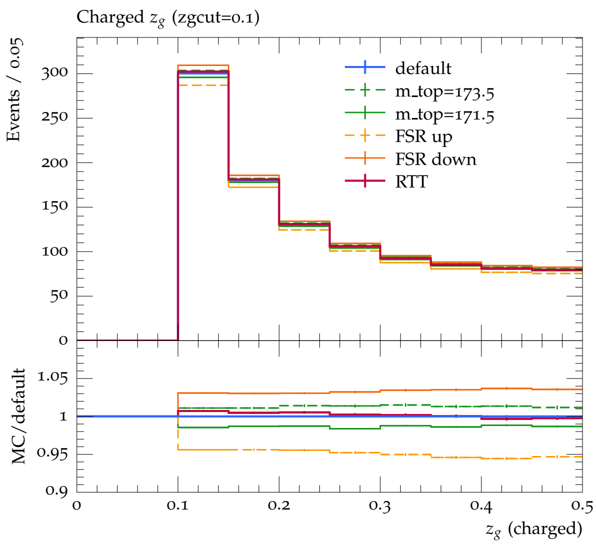

Figure 11: Momentum fraction z_g (charged constituents)

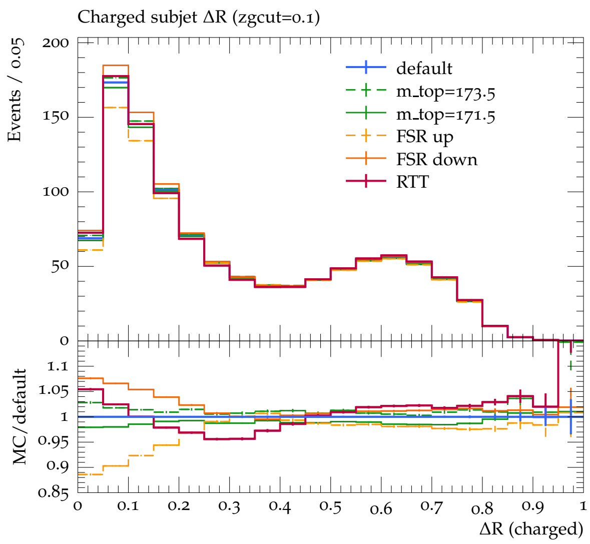

Figure 12: Subjet ΔR (charged constituents)

4.1 Fast simulation of detector effects

To examine the possibility of actually measuring some of these observables, a fast simulation of detector effects was implemented. While Rivet does provide some functions for simulating detector effects, such as general and also experiment specific energy and momentum smearing as well as reconstruction efficiency, functions for the CMS detector are often copies of functions for ATLAS, or lack detail, and based on author suggestions, are to be adjusted for each analysis [7]. Because of this, the analysis implements new functions for gaussian energy smearing (based on [8]) as well as charged particle reconstruction efficiency (based on [9]). To better match the energy resolution achievable by the CMS detector, an algorithm for calorimeter clustering was developed. As this clustering was done in the last two weeks of the project, the developed algorithm is considered preliminary and should be improved in case of further use.

As the subjet ΔR observable is the most promising candidate for future studies of hard radiation, the effects of smearing were tested on it and response matrices were made. The collinear radiation acts the same way for both distributions but the wide angle radiation looks smeared for subjet ΔR for all jet constituents, which is expected as there is lower angular resolution applied. The response matrix for charged constituent ΔR (Fig.16) shows a very narrow diagonal distribution, which is to be expected as for the charged particles only a reconstruction efficiency function was applied. For all constituent particles (Fig.15), there is a larger off-diagonal distribution, showing successful smearing of angular components. Overall these graphs show that smearing effects are working, but do not change the distribution shapes too much, indicating that it should be possible to measure this observable by the CMS detector to distinguish between different models.

As the ΔR distribution for all constituents (Fig.13) show way less events, diminished by a factor of 2 from the original, this indicates that there must be a large amount of rejected events at some point in the calorimeter clustering algorithm or in the plotting sequence. This means that for further use this preliminary code must be improved for more realistic results.

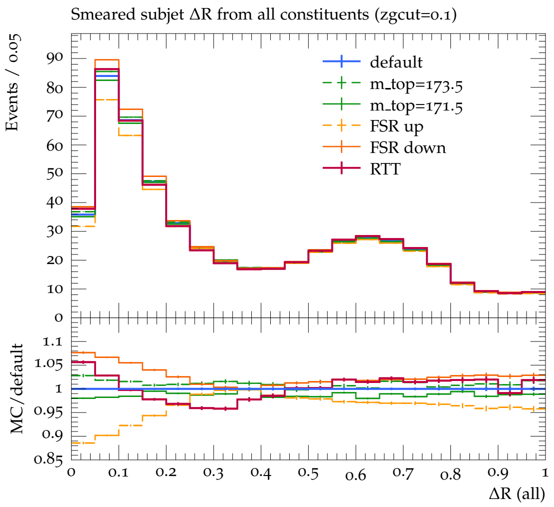

Figure 13: Smeared subjet ΔR (all constituents)

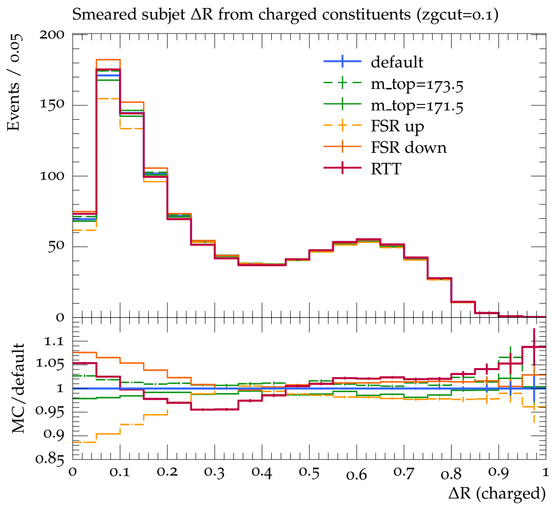

Figure 14: Smeared subjet ΔR (charged constituents)

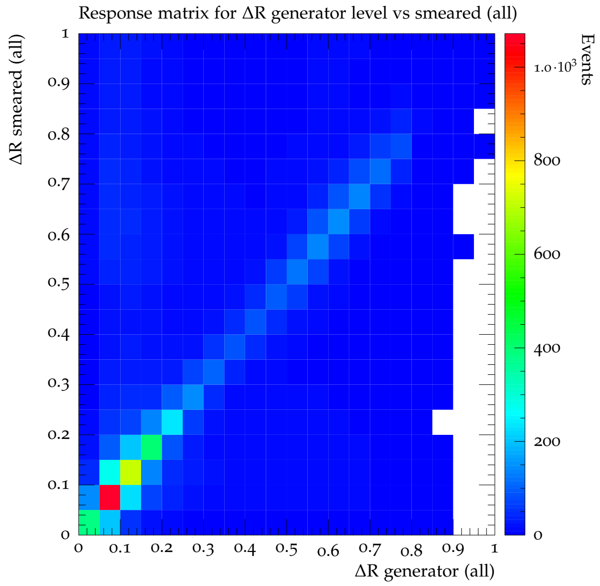

Figure 15: ΔR response matrix (all constituents)

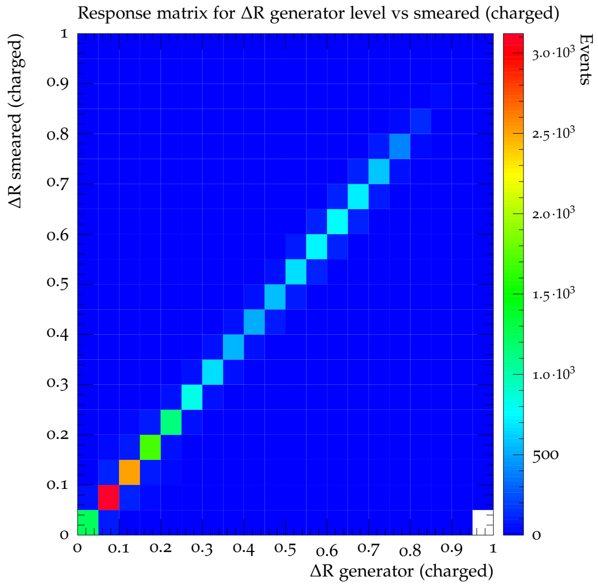

Figure 16: ΔR response matrix (charged constituents)

5 Conclusions

Studies on recoil to top are relevant for improving top quark modelling. The soft jet study shows that for different mass observables, RTT samples were not clearly distinct from FSR variations or top mass variations, and in the 2D plots FSR could not be clearly distinguished from ISR. The fat jet substructure study, particularly the subjet ΔR after soft-drop declustering shows promise in discerning RTT from other variations. The fast detector effect simulation results show promise for the viability of measuring the subjet ΔR in the CMS detector, which would provide more information on the accuracy of RTT in top quark modelling and ultimately lead to a more precise top mass measurement.

References

[1] CMS Collaboration. Measurement of the top quark mass using a profile likelihood approach with the lepton+jets final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C 83 (2023) 963*, February 2023.

[2] P. Skands. Note on RecoilToColored. https://skands.physics.monash.edu/slides/pdf/20-NoteRecoilToColoured.pdf, November 2020.

[3] P. Skands. Notes on Top Quark Modelling in Pythia 8 and Vincia. https://skands.physics.monash.edu/slides/pdf/skands21-top.pdf, March 2021.

[4] V. Slokenbergs M. Seidel, P. Silva. Recoil uncertainty in top mass measurement. https://cds.cern.ch/record/2868479/files/Slokenbergs_summer_report%20(1).pdf, (2023).

[5] Christian Bierlich et al. Robust Independent Validation of Experiment and Theory: Rivet version 3. *SciPost Phys.*, 8:026, 2020.

[6] Marzani S. Soyez G. et al. Larkoski, A.J. Soft Drop. *J. High Energ. Phys. 2014, 146*, May 2014.

[7] Karl Nordström Andy Buckley, Deepak Kar. Fast simulation of detector effects in rivet. *SciPost Phys. 8, 025*, (2020).

[8] CMS Collaboration. The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c. J. Phys.: Conf. Ser. 160 012056, (2009).

[9] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation, Volume 9*, October 2014.