Published for SISSA by Springer

Received: June 27, 2019
Accepted: August 10, 2019
Published: August 30, 2019

Pileup and underlying event mitigation with iterative constituent subtraction

P. Berta,[a,1] **L. Masetti,**[a] **D.W. Miller**[b] **and M. Spousta**[c]

[a] *PRISMA$^+$ Cluster of Excellence and Institute of Physics, Johannes Gutenberg University Mainz, Staudingerweg 7, 55128 Mainz, Germany*

[b] *The Enrico Fermi Institute and the Department of Physics, University of Chicago, 5640 S. Ellis Ave, Chicago, IL, 60637, U.S.A.*

[c] *Institute of Particle and Nuclear Physics, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, 180 00 Prague 8, Czech Republic*

E-mail: peberta@uni-mainz.de, masetti@uni-mainz.de, David.W.Miller@uchicago.edu, Martin.Spousta@mff.cuni.cz

Abstract: The hard-scatter processes in hadronic collisions are often largely contaminated with soft background coming from pileup in proton-proton collisions, or underlying event in heavy-ion collisions. This paper presents a new background subtraction method for jets and event observables (such as missing transverse energy) which is based on the previously published Constituent Subtraction algorithm. The new subtraction method, called Iterative Constituent Subtraction, applies event-wide implementation of Constituent Subtraction iteratively in order to fully equilibrate the background subtraction across the entire event. Besides documenting the new method, we provide guidelines for setting the free parameters of the subtraction algorithm. Using particle-level simulation, we provide a comparison of Iterative Constituent Subtraction with several existing methods from which we conclude that the new method has a significant potential to improve the background mitigation in both proton-proton and heavy-ion collisions.

Keywords: Jet substructure, Hadron-Hadron scattering (experiments), Hard scattering, Jets, Minimum bias

ArXiv ePrint: 1905.03470

[1] Corresponding author.

https://doi.org/10.1007/JHEP08(2019)175

JHEP08(2019)175

Contents

1 Introduction

Precision tests of Standard Model of particle physics as well as searches for new physics at the Large Hadron Collider (LHC) require maximizing the collected data which is primarily achieved by increasing the beam intensities. The dilemma is that high intensities also increase the number of simultaneous proton-proton (pp) interactions in a single colliding bunch crossing of the two beams. Such multiple pp collisions, also known as *pileup*, are then read out as one single event. Pileup contributes a primarily low transverse momentum (p_{T}) and largely diffuse background of particles to the hard scattering process of interest that subsequently distorts many hadronic final state observables, in particular the kinematics and substructures of hadronic jets.

JHEP08(2019)175

In the last part of LHC Run 2, the ATLAS and CMS experiments achieved an average and peak pileup of ∼35 and ∼70 pp simultaneous collisions, respectively [1, 2]. Each additional pileup pp collision at 13 TeV adds an average p_{T} of approximately ∼900 MeV per unit area in the rapidity-azimuth ($y-\phi$) plane. In the upcoming LHC Run 3, an average pileup of 70 collisions per bunch crossing is expected, corresponding to the maximum pileup level of Run 2. The high-luminosity LHC is expected to deliver an average pileup of 200 pp collisions per bunch crossing [3]. An even more intense environment is present in heavy-ion collisions at the LHC and at the Relativistic Heavy Ion Collider (RHIC) where a large underlying event (UE) may lead to a background p_{T} of more than 300 GeV per unit area in $y-\phi$ space [4]. Efficient techniques need to be developed and tested in order to mitigate the impact of these large backgrounds on jet kinematics, jet substructure, and missing E_{T} measurements. Throughout this paper the word *background* will be used to refer either to pileup in pp collisions or to the UE in heavy-ion collisions.

One category of background mitigation techniques includes subtraction methods or algorithms that do not alter the jet definition itself. Examples of algorithms in this category include the extensively used Area Subtraction [5, 6] approach to correct jet kinematics, and its extension to corrections for jet shape observables, the *Shape-expansion* method [7]. Also in this category are numerous algorithms aimed at correcting the inputs to jet reconstruction prior to the application of a jet algorithm: Constituent Subtraction (CS) [8], SoftKiller [9], PUPPI [10], and jet cleansing [11] are among the most widely studied. A second category of background mitigation techniques consists of the so-called *jet grooming* methods, which are often used in the context of highly Lorentz-boosted massive objects, to both improve the precision of the reconstruction itself and mitigate the effects of pileup and UE. Grooming algorithms fundamentally modify the jet definition to be less sensitive to these backgrounds. The most frequently used methods are filtering [12], trimming [13], pruning [14], and soft drop methods [15–17], each of which have configurable parameters that determine their properties as *groomers*. These methods often work with subjets that are determined by the Cambridge-Aachen algorithm [18, 19] or the k_t algorithm [20, 21] applied to a large-radius anti-k_t [22] jet. Specific algorithms or criteria are then applied to these subjets to eliminate soft and/or wide-angle contributions to a jet, which are the most likely to be contaminated by backgrounds. Besides subtraction and grooming methods, techniques based on machine learning have also been implemented and tested recently [23–25]. The potential of improving background mitigation by incorporating charged-track information to subtract neutral pileup is discussed in ref. [26]. Recently, a novel approach to estimate the background is proposed in ref. [27], which aims at reducing the impact of fluctuations in the background on the jet observables and which may be potentially combined with CS algorithm or other subtraction method. A detailed overview of background mitigation techniques can be found in ref. [28].

This paper presents a new background subtraction method for jets, jet substructure observables, and global event observables such as missing E_{T}, which is based on the CS method. In contrast to other background subtraction methods, such as the Area Subtraction or the shape-expansion method, CS corrects for background at the level of jet constituents, which may be particles, tracks from an inner detector, or calorimeter clus-

JHEP08(2019)175

ters [29–31]). Jet kinematics and substructure are simultaneously corrected with this approach. The CS method was successfully used in several measurements at the LHC. The ALICE and CMS experiments use CS in measurements of jet shapes and mass in heavy-ion collisions [32–34], furthermore CMS used CS in the measurement of splitting functions [35]. The CS method has been applied in several performance studies by CMS [36, 37] and ATLAS [38–40] and it is being discussed and tested in the context of future experiments, in particular the Compact Linear Collider (CLIC) and Future Circular Collider (FCC) [41, 42]. The CS method was also used in a recent measurement of jet substructure by STAR at RHIC [43]. Phenomenological studies also consider the CS method, in particular in the context of tagging boosted bosons and top quarks [44–46], searches for new physics [47–50], and structure of parton shower [51]. The new method for background subtraction presented in this paper may help to improve the precision of measurements of Standard Model processes and may make experimental studies less susceptible to increasing backgrounds at colliders. This new method is applicable for the subtraction of both pileup in pp collisions and UE in heavy-ion collisions. In this paper, we test the methods explicitly only in the environment of pp collisions with pileup.

This paper is organized as follows. First, in section 2, an event-wide background subtraction performed using CS is discussed, which was briefly mentioned in ref. [8] as a possible extension of the original CS method. In section 3, the new method for background subtraction is introduced. Then, in section 4, performance of various methods for background subtraction including the new method is presented in the context of jet reconstruction and substructure observables.

2 Event-wide pileup mitigation with CS

The CS algorithm described in ref. [8] corrects individual jets which were already clustered using a certain jet algorithm. We refer to this approach as *Jet-by-jet CS*. However, the jet clustering can be biased by the presence of background resulting in different particle content of jets (this is referred to as "back reaction" in ref. [6]). As briefly mentioned in ref. [8], the CS algorithm can be extended to correct the event constituents before jet clustering. The jets resulting from this *Event-wide CS* correction followed by jet clustering can be more precise than the individually corrected jets. In the following, the Event-wide CS is described assuming massless inputs. A discussion of the correction of massive inputs can be found in appendix C.

The basic ingredient of CS is the background p_{T} density, ρ, which was introduced in the Area Subtraction. Several methods to estimate this quantity are described in [52]. In general, ρ can be estimated as a function of other variables, most commonly as a function of rapidity. The estimated ρ is then used to scale the p_{T} of the *ghosts* in the Event-wide CS. The ghosts are infinitly soft particles (in practice $p_{\mathrm{T}} \approx 10^{-100}\,\mathrm{GeV}$) incorporated into the event such that they uniformly cover the $y-\phi$ plane with high density. Each of these ghosts is massless and covers a fixed area, A_g, in the $y-\phi$ plane. Historically, the ghosts can be used to define the jet area [5, 6] for the Area Subtraction method or perform

JHEP08(2019)175

background subtraction in the Shape-expansion and Jet-by-jet CS methods. In all these methods, their property of infinite softness is essential to not modify the jet clustering sequence. However, for the Event-wide CS, this property is irrevelant and each ghost p_{T} is directly set to $p_{\mathrm{T}}^{g} \equiv A_g \cdot \rho$. Then such ghosts already represent the expected background contribution in the given event, and can be used to correct particles via the Event-wide CS as follows.

Only particles and ghosts with pseudo-rapidity, η, fulfilling $|\eta| < \eta^{\max}$ are used in the correction procedure. The parameter $\eta^{\max}$ defines the detector acceptance of particles.[1] For each pair of particle i and ghost k, a matching scheme is implemented using the distance measure, $D_{i,k}$, defined as

$$D_{i,k} = p_{\mathrm{T},i}^{\alpha} \cdot \Delta R_{i,k}, \tag{2.1}$$

where α is a free parameter and $\Delta R_{i,k}$ is defined as

$$\Delta R_{i,k} = \sqrt{\left(y_i - y_k^g\right)^2 + \left(\phi_i - \phi_k^g\right)^2}. \tag{2.2}$$

The list of all distance measures, $\{D_{i,k}\}$, is sorted from the lowest to the highest values. The background removal proceeds iteratively, starting from the particle-ghost pair with the lowest $D_{i,k}$. At each step, the momentum p_{T} of each particle i and ghost k are modified as follows.

$$\begin{aligned} \text{If } p_{\mathrm{T},i} \geq p_{\mathrm{T}k}^{g}: \quad & p_{\mathrm{T},i} \longrightarrow p_{\mathrm{T},i} - p_{\mathrm{T}k}^{g}, \\ & p_{\mathrm{T}k}^{g} \longrightarrow 0; \\ \text{otherwise:} \quad & p_{\mathrm{T},i} \longrightarrow 0, \\ & p_{\mathrm{T}k}^{g} \longrightarrow p_{\mathrm{T}k}^{g} - p_{\mathrm{T},i}. \end{aligned} \tag{2.3}$$

The iterative process is terminated when $\Delta R_{i,k} > \Delta R^{\max}$ where $\Delta R^{\max}$ is a free parameter.[2] The output of the correction procedure is a set of 4-momenta representing the background-corrected event. Any operation can be done on these output particles — most commonly jet clustering, evaluation of global event shapes or missing transverse energy. Besides providing the ability to correct the full event, which is documented for the case of missing transverse energy in section 4.5, the performance of the background subtraction is also improved with respect to the original CS as documented in section 4. A detailed discussion of the choice of the free CS parameters is provided in appendix A.

[1] In the software implementation of the Event-wide CS algorithm in `FastJet Contrib` [53], the user can define arbitrary phase-space using the `fastjet::Selector` class from `FastJet` [52, 54]. In that way, phase space asymmetric in η can be used or the correction can be done only for particles below certain p_{T} threshold.

[2] The meaning of the $\Delta R^{\max}$ parameter was different in the original description of the CS procedure in ref. [8]. In the software implementation in `FastJet Contrib`, the meaning described in this paper is used since version 1.022.

JHEP08(2019)175

3 Iterative Constituent Subtraction

Iterative Constituent Subtraction (ICS) is a new background mitigation method that extends the concept of the original CS method. ICS applies the event-wide implementation of CS iteratively with finite $\Delta R^{\max}$ value. After each iteration, any remaining unsubtracted background estimate is redistributed uniformly across the entire event and another CS procedure is performed. The exact algorithmic procedure using N^{iter} iterations is the following:

1. Estimate ρ using a given estimation method and create ghosts with $p_{\text{T}}^{g} = \rho A_g$. The ρ can be a function of other variables.

2. Perform Event-wide CS correction using a finite $\Delta R^{\max}$ value. Define *input ghosts* as ghosts present in the event before this CS correction. Define *output ghosts* as ghosts remaining in the event after this CS correction.

3. Compute the scalar p_{T} sum of the input ghosts, $p_{\text{T}}^{\text{input}}$, and of the output ghosts, $p_{\text{T}}^{\text{output}}$. Update the input ghosts by scaling their p_{T} by a factor $p_{\text{T}}^{\text{output}}/p_{\text{T}}^{\text{input}}$.

4. Perform next iteration by going back to step 2 using the updated input ghosts. The parameters of CS, such as $\Delta R^{\max}$, may be changed for each iteration.

The step 4 is performed $(N^{\text{iter}} - 1)$-times. The algorithm with $N^{\text{iter}} = 2$ is illustrated on an example event in figure 1.

An option exists for the above algorithm that avoids the usage of ghosts in the second or higher iteration which were not fully subtracted in the previous iteration. With this option, the $p_{\text{T}}^{\text{input}}$ is evaluated using only ghosts which were fully subtracted in the previous iteration, while the unsubtracted ghosts are discarded for the actual iteration. This approach avoids placing the expected background deposition to positions where there is small chance for combining ghosts with real particles in the second or higher iteration. We refer to this option as *ghost removal*. The impact of this option on the performance depends heavily on the chosen $\Delta R^{\max}$ parameters for each iteration.

In our performance studies, it was found that ICS outperforms both Jet-by-jet CS and Event-wide CS (described in section 2) as documented in section 4. A detailed discussion of the choice of the configuration of the ICS method is provided in appendix B. The ICS method is implemented in `FastJet Contrib` [53] since version 1.038.

4 Performance

In this section, we discuss the performance of new methods presented in sections 2 and 3. The definitions of jet shapes which are commonly used in jet substructure studies are summarized in section 4.1. In section 4.2, the metrics used to measure the performance of background subtraction methods is discussed. The samples and general settings used for the performance evaluation are described in section 4.3. The other sections contain several performance studies. Studies for ICS, Event-wide CS, Jet-by-jet CS and Area Subtraction

JHEP08(2019)175

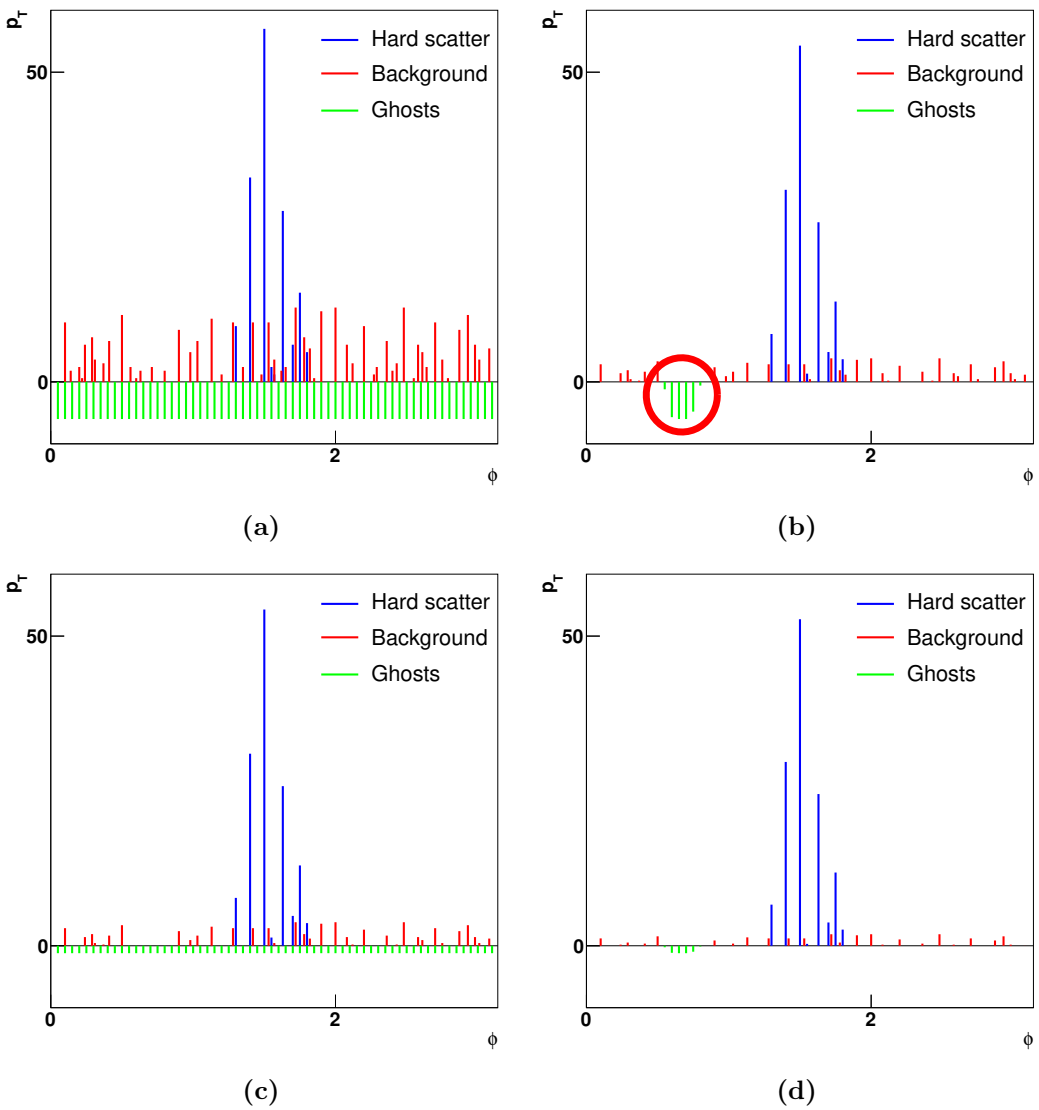


Figure 1. Illustration of the ICS method with 2 iterations assuming only one dimension in azimuth. The event before the first iteration (1a) contains hard-scatter particles, background particles and ghosts with p_{T} corresponding to ρ in that event. To emphasize the fact that the ghosts are subtracted, their p_{T} is negative in these figures. After the first iteration with finite ΔR^{max} (1b), ghosts with unsubtracted p_{T} can remain in the event (emphasized by a red circle). The scalar p_{T} sum of the unsubtracted ghosts is redistributed uniformly in the $y-\phi$ plane (1c) or according to the actual ρ dependence. After the second iteration (1d), the unsubtracted background p_{T} is strongly reduced.

for jets are presented in section 4.4. Studies for missing transverse energy are presented in section 4.5. An extensive comparison among a several selected methods using a common open-source software framework for background subtraction performance comparisons is provided in section 4.6.

JHEP08(2019)175

4.1 Jet shape definitions

While kinematic properties of jets can be uniquely characterized by four-momentum, $P^\mu = (p_\mathrm{T}, \eta, \phi, m)$, the internal structure of jet cannot be characterized by a single observable. Instead, various observables have been proposed in order to capture many internal properties of jets. These different observables are generically referred to as jet shapes. This term extends the meaning of the observable $\rho(r)$, originally called jet shape, which was measured at various experiments [55–61] and which quantifies the radial energy flow, from the jet axis $r = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. This quantity, along with other observables such as fragmentation function, can be used to quantify the internal structure of QCD jets in the context of the parton shower evolution. More recently, jet shape observables have been extensively employed as tools in measurements of Lorentz-boosted massive objects [62–65]. In such cases, jets have fundamentally different internal structures as compared to jets formed from light quarks and gluons. One canonical example is the two-prong structure formed from a collimated pair of heavy quarks from the decay of a boosted Higgs boson [12]. Various jet shape and structure observables can be used to distinguish jets formed by massive boosted object decays from jets formed solely by parton showering and hadronization of light quarks and gluons. In this study, we use three representative jet shape observables — *jet width*, two-to-one and three-to-two ratios of *N-subjettiness* (τ_{21} and τ_{32}) — and the *jet mass*.

Jet mass, $m = \sqrt{P^\mu P_\mu}$, is a basic observable used to identify boosted hadronically decaying objects such as the $W^\pm$ boson (see e.g. ref. [66]). Jet width (also known as *linear radial moment* or *girth*) is defined as the first moment of the radial flow of transverse momentum or energy,

$$\text{jet width} = \frac{\sum_i p_{\mathrm{T},i} \Delta R_i}{\sum_i p_{\mathrm{T},i}}, \tag{4.1}$$

where $p_{\mathrm{T},i}$ is magnitude of the transverse momentum of jet constituent i and ΔR_i is the distance between jet constituent i and the jet axis. Jet width has been shown to provide e.g. a good discriminating power between quark-initiated and gluon-initiated jets [67].

The τ_{21} and τ_{32} are two-to-one and three-to-two ratio of N-subjettiness, respectively. Subjettiness was introduced to provide an enhanced discriminating power for identifying boosted massive objects [68, 69]. The N-subjettiness is defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{\mathrm{T}k} \cdot \min(\Delta R_{1k}, \Delta R_{2k}, \ldots, \Delta R_{Nk}) \ , \quad \text{with,} \quad d_0 \equiv \sum_k p_{\mathrm{T}k} \cdot R \tag{4.2}$$

where R is the distance parameter of the jet algorithm, $p_{\mathrm{T}k}$ is the transverse momentum of constituent k and ΔR_{ik} is the distance between a subjet i and a constituent k. The N subjets are defined by re-clustering the constituents of the jet with exclusive version of the k_t algorithm and requiring that exactly N subjets are found.

4.2 Quantifying performance of background subtraction

The following figures of merit are used to quantify the performance of jet reconstruction [70, 71]: jet energy scale (JES), jet energy resolution (JER), jet reconstruction efficiency, and rate of fake jets. JES is also sometimes called linearity. These quantities

JHEP08(2019)175

are calculated using Monte Carlo generators coupled with full detector simulations for a particular experiment by comparing particle-level jets (so-called *true jets*) with jets reconstructed using the outputs of the detector simulation. JES and JER characterize the mean and root-mean-square, respectively, of the difference between the transverse momentum of particle level jet, $p_{\mathrm{T}}^{\mathrm{true}}$, and the transverse momentum of jet reconstructed in the detector, $p_{\mathrm{T}}^{\mathrm{reco}}$. More explicitly, JES $= \left\langle p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{true}} \right\rangle$ and JER $= \sigma(p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{true}})$. The JES is determined by a combination of the response of the detector and the performance of algorithms used for the jet calibration and background subtraction. The JER of calorimeter jets can be factorized as follows

$$\sigma\left(\frac{p_{\mathrm{T}}^{\mathrm{reco}}}{p_{\mathrm{T}}^{\mathrm{true}}}\right) = \frac{a}{\sqrt{p_{\mathrm{T}}^{\mathrm{true}}}} \oplus \frac{b}{p_{\mathrm{T}}^{\mathrm{true}}} \oplus c, \tag{4.3}$$

where a is the stochastic term, b is the noise term and c is the constant term [72, 73]. The stochastic and constant terms are governed by the response of the detector to the particle shower. The noise term is largely determined by fluctuations of backgrounds, which may include electronic noise, pileup, or the underlying event. At the LHC, the contributions due to pileup tend to dominate. In the case of an average background subtraction such as the Area Subtraction, the noise term is given by the root-mean-square of p_{T} evaluated in the area of a jet excluding the jet signal, $b = \mathrm{RMS}(p_{\mathrm{T}}^{\mathrm{area}})$. The stochastic and constant terms are not significantly affected by the subtraction algorithm [72], but they can be reduced by using the information about the jet internal structure in the calibration procedure [74] or by combining the information from calorimeter with the information from tracking [75]. The noise term cannot be reduced by the calibration procedure, but it can be reduced by subtraction procedure or using some additional information about the background compared to the basic information about its average density. This can be, for example, the information about pointing of jet constituents to the primary vertex (e.g. *charged hadron subtraction* used by CMS [76] or *jet-vertex association* used by ATLAS [77]) or by noise suppression at the sub-constituent level of calorimeter cells [78, 79].

Jet reconstruction efficiency quantifies the probability that a jet is found given the presence of a true jet, subject to specific kinematic constraints for a particular efficiency evaluation. As such, the jet reconstruction efficiency is indirectly affected by both the JES and the JER. The rate of fake jets is important in the case of large backgrounds, such as in heavy-ion collisions, where correlated background fluctuations can lead to a large rate of fake jets. Consequently, this rate can be reduced by reducing the fluctuations in the background for which the JER is the relevant performance metric. We therefore focus on the JES and JER in order to characterize the jet reconstruction performance.

The JES and JER can be generalized to any of the observable quantities that characterize the jet kinematics or shapes discussed in section 4.1. We define the *bias* and *resolution* of a given observable, x, as follows:

$$\text{bias} = \frac{\langle x^{\mathrm{rec}} - x^{\mathrm{true}} \rangle}{\langle x^{\mathrm{true}} \rangle}, \quad \text{resolution} = \frac{\mathrm{RMS}(x^{\mathrm{rec}} - x^{\mathrm{true}})}{\langle x^{\mathrm{true}} \rangle}. \tag{4.4}$$

The difference between x^{rec} and x^{true} is used instead of their ratio in order to avoid excessive values of bias for small values of x^{true}. The denominator in the bias and resolution

JHEP08(2019)175

allows for an easier comparison among different quantities. Similarly to the case of the JES and JER, these quantities characterize the primary aspects of the performance of the background subtraction. These quantities are therefore evaluated simultaneously for all the jet shape observables discussed in section 4.1.

4.3 Test samples and configuration of subtraction

The performance of subtraction methods is evaluated in simulated pp collisions at $\sqrt{13}$ TeV using events with boosted top quarks from the decay $Z' \to t\bar{t}$ of a hypothetical boson Z' with a mass of 1.5 TeV. These simulated hard-process events are referred to as *true* events. To obtain the *reconstructed* events with pileup included, the true events are overlaid with inclusive pp collision events that represent pileup. The number of overlaid inclusive events, N_{PU}, has a uniform distribution from 0 to 140. All event generation is performed with PYTHIA 8.180 [80, 81] using the tune 4C and CTEQ 5L parton density functions [82]. The hard process is generated without underlying event.

A pseudo-detector simulation is then used. All particles are grouped into towers of size 0.1×0.1 in the pseudorapidity-azimuth $(\eta - \phi)$ space. The tower energy is obtained as the sum of energies of particles pointing to that tower. All neutrinos and muons are discarded. Only towers with $|\eta| < 4.0$ are selected. The mass of each tower is set to 0. The η and ϕ of each tower is randomly smeared using a Gaussian kernel with width of 0.1 (but maximally up to 0.2).

The reconstructed events are corrected with both ICS and multiple other pileup mitigation techniques. To evaluate the performance of the various methods, the jets from the corrected events are compared to jets from the true events using the quantities defined in section 4.2. To factorize the effect of pileup, the same detector simulation is used for both, true and reconstructed, events. Two jet definitions are used: anti-k_t algorithm with the distance parameter 0.4 and 1.0. Only true jets with $|\eta| < 3$ and $p_{\mathrm{T}} > 20$ GeV are used.

All jet finding and background estimation is performed using `FastJet` 3.3.1 [52, 54]. The event energy density ρ is estimated as a function of y using the `GridMedianBackgroundEstimator` tool from `FastJet` with a grid spacing of 0.5 using the particles up to $|y| = 4$. Since the inputs are massless, there is no need to derive the background estimate for the mass term of pileup, ρ_m. The same ρ estimation is used for the Area Subtraction and the CS-based methods. The Area Subtraction method was carried out as 4-vector subtraction using the `fastjet::Subtractor` class from `FastJet` 3.3.1. Unphysical situations with negative corrected mass are avoided by enabling the `safe_mass` option.

For the CS-based methods, we used `FastJet Contrib` version 1.042. The choice of the optimal CS parameters and ICS configurations is discussed in detail in appendices A and B. The following configurations are found to be optimal for the majority of studies presented in sections 4.4 and 4.5. The Jet-by-jet CS uses parameters $\Delta R^{\max} = \infty$, $\alpha = 0$, and $A_g = 0.0025$ for both jet definitions. The Event-wide CS uses parameters $\alpha = 1$, $A_g = 0.0025$, and $\Delta R^{\max}$ parameters $\Delta R^{\max} = 0.25$ and $\Delta R^{\max} = 0.7$ for anti-k_t $R = 0.4$ and $R = 1.0$ jet definitions, respectively. The ICS method is used with two iterations using the ghost removal option. For both iterations, the same α and A_g parameters are used, $\alpha = 1$ and $A_g = 0.0025$, while the $\Delta R^{\max}$ is different for each iteration and also depends on the jet

JHEP08(2019)175

definition. For anti-k_t $R = 0.4$ jets, parameters $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.1$ are used for the first and second iteration, respectively. For anti-k_t $R = 1.0$ jets, the parameters $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.35$ are used.

4.4 Performance for jet kinematics and substructure

The performance of the ICS method applied to jets is evaluated for both jet kinematic and jet shape observables, which are defined in section 4.1. The bias and resolution (defined in section 4.2) for the observables are studied as a function of number of pileup interactions (N_{PU}) and jet p_{T}, as well as for two choices of anti-k_t distance parameter, $R = 0.4$ and $R = 1.0$. The performance of ICS is compared to the performance of the Area Subtraction, Jet-by-jet CS (both introduced in section 1), and the Event-wide CS (discussed in section 2) using the configurations described in section 4.3.

Perhaps the most illustrative demonstration of the efficiency of any pileup correction is the extent to which a given algorithm is able to reduce the dependence of an observable on the amount of pileup in the event, as parameterized by N_{PU}. Figures 2 and 3 show the impact of pileup on the p_{T} and mass, respectively, for large-radius ($R = 1.0$) jets containing the decay products of boosted top-quarks from $Z' \to t\bar{t}$ (see section 4.3). Only a narrow true jet $p_{\mathrm{T}}^{\mathrm{true}}$ range is chosen for these figures, $250\,\mathrm{GeV} \leq p_{\mathrm{T}}^{\mathrm{true}} < 300\,\mathrm{GeV}$. The four correction algorithms considered for comparison demonstrate the improvements achievable in terms of the bias (figures 2a and 3a) and resolution (figures 2b and 3b) in each case, as a function of N_{PU}.

Each of the algorithms considered is able to remove the bias introduced by pileup in both the p_{T} and the mass of these jets to approximately the same degree. However, the precision of these corrections, as quantified by the resolution of the corrected measure, can vary significantly. The Event-wide CS and the ICS corrections are both observed to improve the resolution of the jet p_{T} and mass by up to 30% at large N_{PU}, as compared to the Jet-by-jet CS correction. Across the full range of N_{PU} considered, ICS exhibits an improvement beyond that of the CS correction alone by approximately 5–10% in the resolution of both the jet p_{T} and the jet mass.

The ability of each algorithm to mitigate the impacts of pileup depends on more than just the amount of pileup considered. As mentioned earlier, the size of the jet and the kinematic range (e.g. high or low true $p_{\mathrm{T}}^{\mathrm{true}}$) can affect the performance significantly. Moreover, the impact of pileup on certain observables can be larger, thus reducing the effectiveness of certain approaches. Figures 4 and 5 summarize the results of a comprehensive study of the performance of each of the four algorithms under consideration. Results are reported in a narrow range of high pileup, $N_{\mathrm{PU}} = 100$–120, for jets from the $Z' \to t\bar{t}$ process, for the following:

- **jet radius:** $R = 0.4, 1.0$
- **true jet p_{T}:**
  - 7 bins of $p_{\mathrm{T}}^{\mathrm{true}}$ in the range $p_{\mathrm{T}}^{\mathrm{true}} = 20$–$600\,\mathrm{GeV}$ for $R = 0.4$
  - 4 bins of $p_{\mathrm{T}}^{\mathrm{true}}$ in the range $p_{\mathrm{T}}^{\mathrm{true}} = 200$–$600\,\mathrm{GeV}$ for $R = 1.0$

JHEP08(2019)175

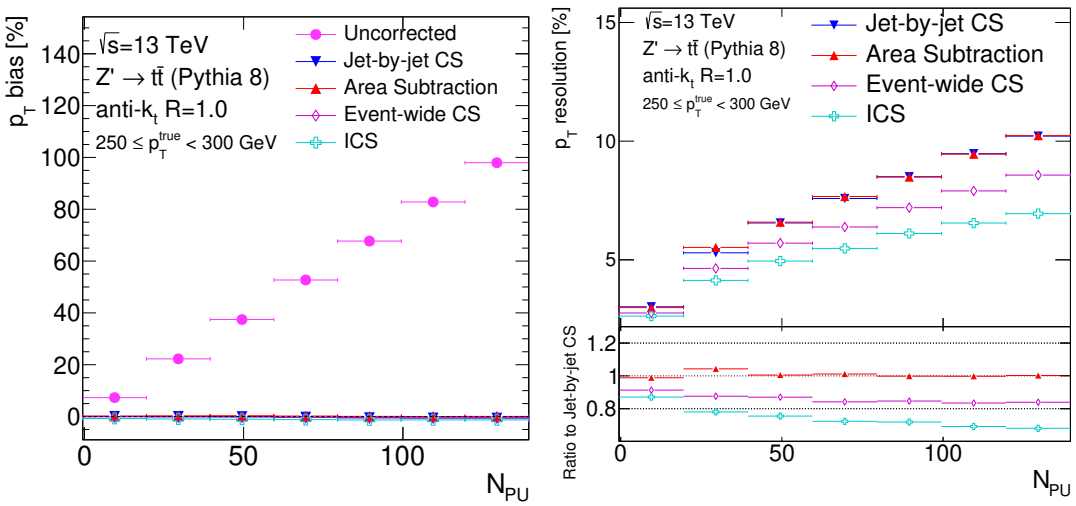


Figure 2. Dependence of jet p_{T} bias (left) and resolution (right) on N_{PU} for four pileup correction methods (Jet-by-jet CS, Area Subtraction, Event-wide CS, and ICS).

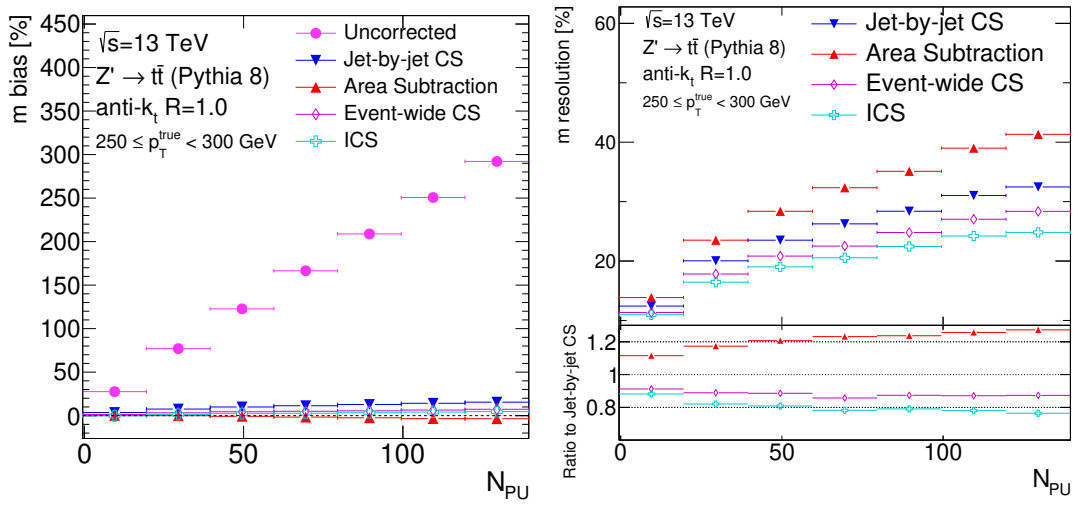


Figure 3. Dependence of jet mass bias (left) and resolution (right) on N_{PU} for four pileup correction methods (Jet-by-jet CS, Area Subtraction, Event-wide CS, and ICS).

- **observable:**
  - η, mass, p_{T}, width for both $R = 0.4, 1.0$
  - τ_{21} and τ_{32} for only $R = 1.0$

The outcome is a set of 52 comparisons each for the bias and resolution after pileup correction.

JHEP08(2019)175

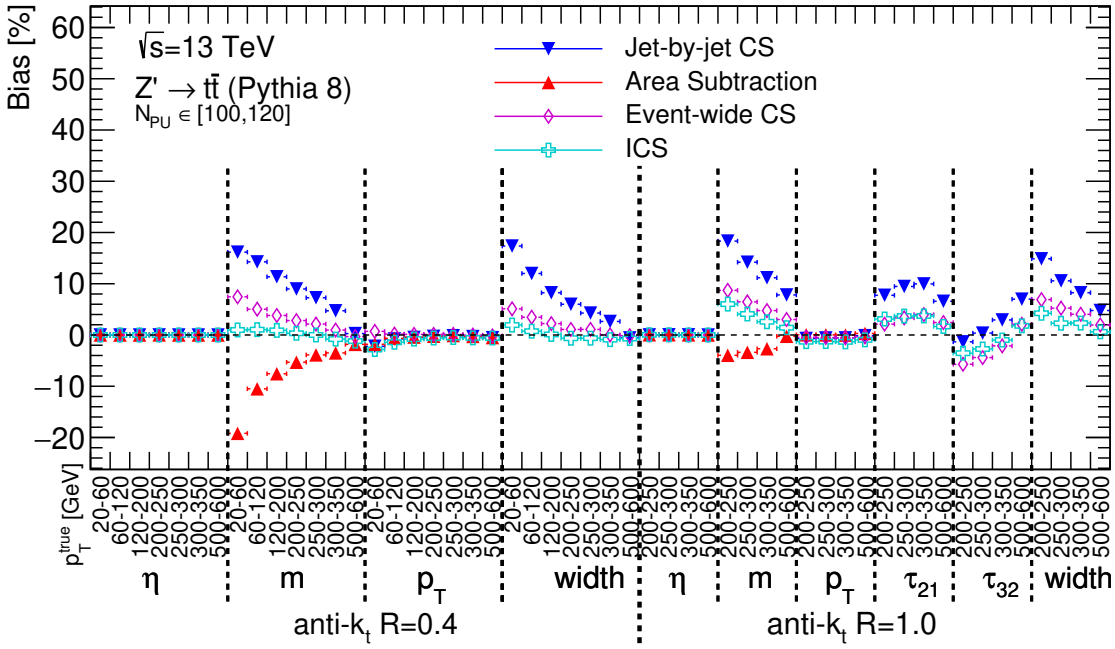


Figure 4. Performance of the pileup subtraction evaluated in terms of the bias for different observables: jet η, mass, p_{T}, width for $R = 0.4$ jets and jet η, mass, p_{T}, width, τ_{21}, τ_{23} for $R = 1.0$ jets. Each bin on the x-axis represents a given range of true jet $p_{\mathrm{T}}^{\mathrm{true}}$ defined by the bin-label. Bins on the x-axis are grouped to distinguish observables and jet radii. Four algorithms are compared: Jet-by-jet CS, Area Subtraction, Event-wide CS, and ICS algorithm.

The results of these comparisons provide a rich set of information from which a few conclusions may be reliably drawn. The ability of algorithms to remove the bias is practically identical in the case of jet p_{T} and jet η. The ability to correct the bias is significantly improved for the Event-wide CS and ICS algorithms compared to Jet-by-jet CS algorithm in the case of jet mass, jet width, and τ_{21}. For τ_{32}, the performance of Jet-by-jet CS, Event-wide CS and ICS algorithms is very similar. The bias generally decreases with increasing jet $p_{\mathrm{T}}^{\mathrm{true}}$ and tend to converge to zero for all the algorithms, both jet radii, and almost all the observables. Such a convergence in the behavior of algorithms is however not seen in the case of resolution, where significant differences among algorithms persist in the full kinematic range and in some cases even tend to get larger at higher $p_{\mathrm{T}}^{\mathrm{true}}$.

A clear improvement in the resolution is seen for Event-wide CS and ICS algorithms compared to both Jet-by-jet CS algorithm and Area Subtraction. While the resolution from ICS and Event-wide CS is practically the same for $R = 0.4$ jets, a significant difference is seen for $R = 1.0$ jets where ICS algorithm outperforms the Event-wide CS. These observations hold for all observables studied and for all the jet $p_{\mathrm{T}}^{\mathrm{true}}$ bins. This conclusion together with the conclusion on the bias implies that ICS algorithm provides the largest and most consistent improvements in the performance out the four algorithms tested.

JHEP08(2019)175

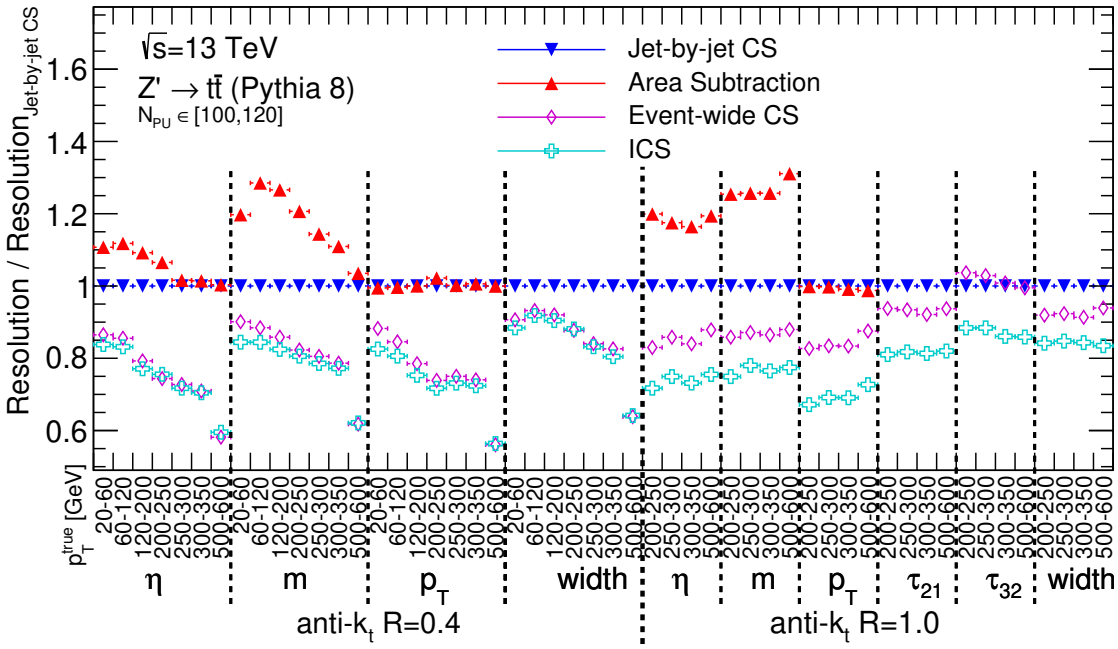


Figure 5. Performance of the pileup subtraction evaluated in terms of the ratio of resolution with respect to the resolution from Jet-by-jet CS algorithm for different observables: jet η, mass, p_{T}, width for $R = 0.4$ jets and jet η, mass, p_{T}, width, τ_{21}, τ_{23} for $R = 1.0$ jets. Each bin on the x-axis represents a given range of true jet $p_{\mathrm{T}}^{\mathrm{true}}$ defined by the bin-label. Bins on the x-axis are grouped to distinguish observables and jet radii. Four algorithms are compared: Jet-by-jet CS, Area Subtraction, Event-wide CS, and ICS algorithm.

4.5 Performance for missing transverse energy

In addition to evaluating the performance of ICS and related background subtraction algorithms in terms of their impacts on jets, we also studied the extent to which improvements might be gained in the measurement of missing transverse energy, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ in the event is defined as a 2-vector calculated as the negative vector sum of the p_{T} of all physics objects in the event. For this study, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is calculated using the vector sum of all stable particles, while jets are not used in the calculation. The results are reported in terms of just one component of the 2-vector, $E_{\mathrm{T},x}^{\mathrm{miss}}$.

Figure 6 shows as a function of N_{PU} a measure of the resolution of the $E_{\mathrm{T},x}^{\mathrm{miss}}$, calculated as the RMS of the difference between the reconstructed and true $E_{\mathrm{T},x}^{\mathrm{miss}}$ in the event. The uncorrected resolution is reported as well as the results of applying four different subtraction algorithms to the entire event: SoftKiller (using grid-size parameter of 0.6), Event-wide CS and ICS with the same configurations as for anti-k_t $R = 0.4$ jets described in section 4.3, and combination of Event-wide CS followed by SoftKiller (tested by the ATLAS Collaboration [40]). The optimal grid-size parameter for SoftKiller depends on the N_{PU}. For example, grid-size parameter of 0.5 provides by ∼5% better resolution for $N_{\mathrm{PU}} < 60$, but by ∼5% worse resolution for $N_{\mathrm{PU}} > 100$ compared to the configuration with the grid-size of 0.6. The grid-size parameter of 0.7 provides by ∼3% better resolution

JHEP08(2019)175

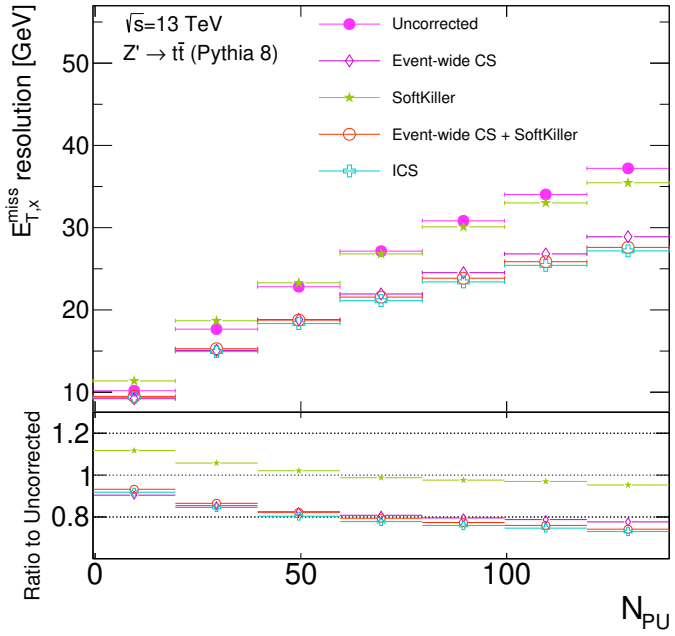


Figure 6. Resolution of the x-component of the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$.

for $N_{\mathrm{PU}} > 100$, but by ∼10% worse resolution for $N_{\mathrm{PU}} < 60$ compared to the configuration with the grid-size of 0.6. For this reason we show SoftKiller with the grid-size parameter of 0.6 which is a good compromise. We found that the best resolution for the CS and SoftKiller combination can be achieved by using CS parameters $\Delta R^{\mathrm{max}} = 0.2$ and $\alpha = 1$, and SoftKiller grid-size parameter of 0.5.

For all the tested methods, the resolution of the $E_{\mathrm{T},x}^{\mathrm{miss}}$ worsens with increasing pileup, as expected. However, the three methods Event-wide CS, ICS, and Event-wide CS followed by SoftKiller reduce this degradation by more than 20% at large N_{PU}, while SoftKiller alone does not perform so well. This along with results presented in the previous section demonstrate the stability and good performance of the ICS method.

4.6 Performance using the framework from the 2014 Pileup Workshop

We compare the performance of the new method with other methods using the common open-source software framework [83] defined at the "Pileup Workshop" held in May 2014 at CERN [84]. The code used to obtain the results in this section is located in the folder `comparisons/ICS` of this framework, version 1.1.0.

The samples used for the comparison presented here are available from ref. [85]. Four hard-scatter physics processes overlaid with a fixed number of pileup events are used. The four hard-scatter samples are dijet events with at least one reconstructed anti-k_t $R = 0.4$ jet satisfying $p_{\mathrm{T}} \geq 20$, 50, 100 or 500 GeV. B-hadrons are kept stable and UE is not simulated. The hard-scatter and pileup samples are simulated using Pythia 8.185 with tune 4C using proton-proton collisions at $\sqrt{s} = 14$ TeV. Four pileup conditions are used: $N_{\mathrm{PU}} = 30$, 60, 100, and 140. The performance is evaluated for anti-k_t $R = 0.4$ jets.

JHEP08(2019)175

The mass of all particles is set to zero, preserving p_{T}, rapidity, and azimuth. Only particles with $|y| < 4$ are used. The corrections are done on events without any detector simulation but assuming idealised tracker, where the information about the origin of charged particles (pileup or hard-scatter event) is known. In that way, the charged pileup particles can be directly removed and/or the information about charged particles can be further used to correct the neutral particles. We compare the CS-based methods with Area Subtraction, SoftKiller, and PUPPI, for which we use the same configurations as for the comparison in ref. [28] except slightly different ρ estimation for the Area Subtraction as described below. Summary of the used configurations and usage of the information about charged particles is the following:

- **Area Subtraction:** all charged pileup particles are discarded. The ρ is estimated using only neutral particles with a grid-size parameter of 0.6 and rapidity rescaling. The protection against negative masses after subtraction is enabled. To determine the jet area, the ghosts are placed up to the edge of the particle rapidity acceptance, $|y| < 4$, with a ghost area of 0.01.

- **SoftKiller:** all charged pileup particles are discarded at the beginning. SoftKiller with grid-size parameter of 0.5 is applied on the neutral particles in the event. The used grid-size parameter is the same as in the `comparisons/review` in ref. [83] version 1.0.0 which was used for the publication [28].[3]

- **PUPPI:** the implementation is taken from the `comparisons/review` in ref. [83] version 1.0.0, which is the original implementation provided by the PUPPI authors in the context of the 2014 Pileup Workshop.

- **Jet-by-jet CS:** all charged pileup particles are discarded. Same ρ estimation as for the Area Subtraction is used. Only neutral particles are corrected using CS with parameters: $\Delta R^{\mathrm{max}} = \infty$, $\alpha = 0$, and $A_g = 0.01$.

- **Event-wide CS:** all charged pileup particles are discarded. Same ρ estimation as for the Area Subtraction is used. Only the neutral particles are corrected using the Event-wide CS with parameters: $\Delta R^{\mathrm{max}} = 0.25$, $\alpha = 1$, and $A_g = 0.005$.

- **CS+SoftKiller:** all charged pileup particles are discarded. First Event-wide CS is applied as described in the previous point (with the same CS parameters). Then the corrected neutral particles are further corrected using SoftKiller (grid-size parameter of 0.6).[4]

- **ICS:** all charged pileup particles are discarded. Same ρ estimation as for the Area Subtraction. Only the neutral particles are corrected using the ICS method with two

[3] We have investigated the performance of other grid-size parameters and confirmed that the value of 0.5 represents the optimal choice.

[4] We have investigated the performance for other CS parameters and SoftKiller grid-size parameter and identified this to be the optimal configuration.

JHEP08(2019)175

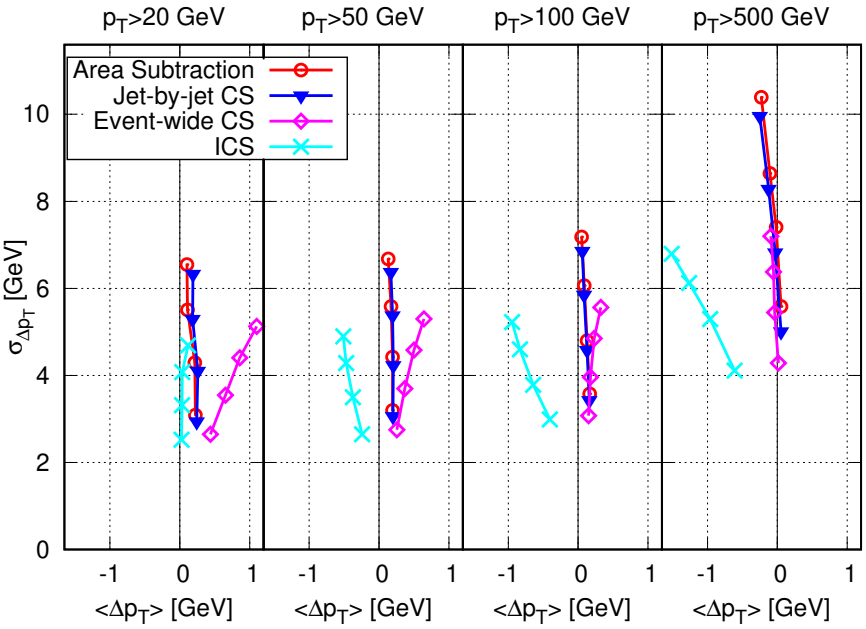


Figure 7. Comparison of the jet p_{T} resolution as a function of jet p_{T} bias for the Area Subtraction, jet-by-jet CS, Event-wide CS and ICS methods. Dijet events are used with different jet p_{T} cut in each panel. Each curve corresponds to a different method and the 4 points on each curve correspond to $N_{\mathrm{PU}} = 30$, 60, 100 and 140 from bottom to top.

iterations and without ghost removal. Parameters: $\alpha = 1$ and $A_g = 0.005$ for both iterations, $\Delta R^{\mathrm{max}} = 0.2$ and $\Delta R^{\mathrm{max}} = 0.15$ for the first and the second iteration, respectively.

`FastJet` v3.3.2 is used for jet clustering, background estimation and Area Subtraction. `FastJet Contrib` v1.038 is used for CS-based methods and SoftKiller. To compare the performance of the methods, the average bias and resolution for the jet p_{T} and jet mass are evaluated. From each event, the two true hardest jets are selected with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|y| < 2.5$. Each corrected jet is matched to the closest true jet requiring $\sqrt{\Delta y^2 + \Delta\phi^2} < 0.3$ (the matching efficiency is above 99.5%). To follow recommendations from the workshop, the average jet p_{T} bias, $\langle \Delta p_{\mathrm{T}} \rangle$, and p_{T} resolution, $\sigma_{\Delta p_{\mathrm{T}}}$, is evaluated as the average and RMS from the p_{T} difference between true and matched corrected jet p_{T}, respectively. Similary, the mass bias, $\langle \Delta m \rangle$, and resolution, $\sigma_{\Delta m}$, are evaluated.

The performance of the Area Subtraction, Jet-by-jet CS, Event-wide CS and ICS methods is shown in figures 7 and 8. Qualitatively, the differences between the individual methods are the same as presented in section 4.4 where, however, a different signal sample and detector simulation are used. Both, Event-wide CS and ICS, improve the resolution significantly with respect to the Area Subtraction and Jet-by-jet CS. The ICS method has slightly better resolution with smaller biases compared to the Event-wide CS.

The performance of the PUPPI, SoftKiller, CS+SoftKiller, and ICS methods is shown in figures 9 and 10. All the compared methods lead to varying levels of bias for some ob-

JHEP08(2019)175

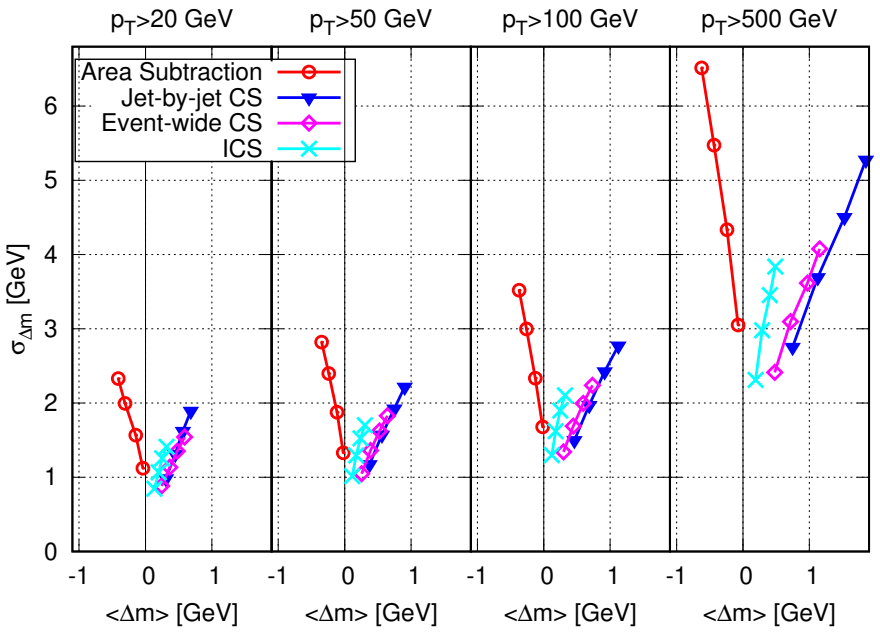


Figure 8. Comparison of the jet mass resolution as a function of jet mass bias for the Area Subtraction, jet-by-jet CS, Event-wide CS and ICS methods. Dijet events are used with different jet p_{T} cut in each panel. Each curve corresponds to a different method and the 4 points on each curve correspond to $N_{\mathrm{PU}} = 30$, 60, 100 and 140 from bottom to top.

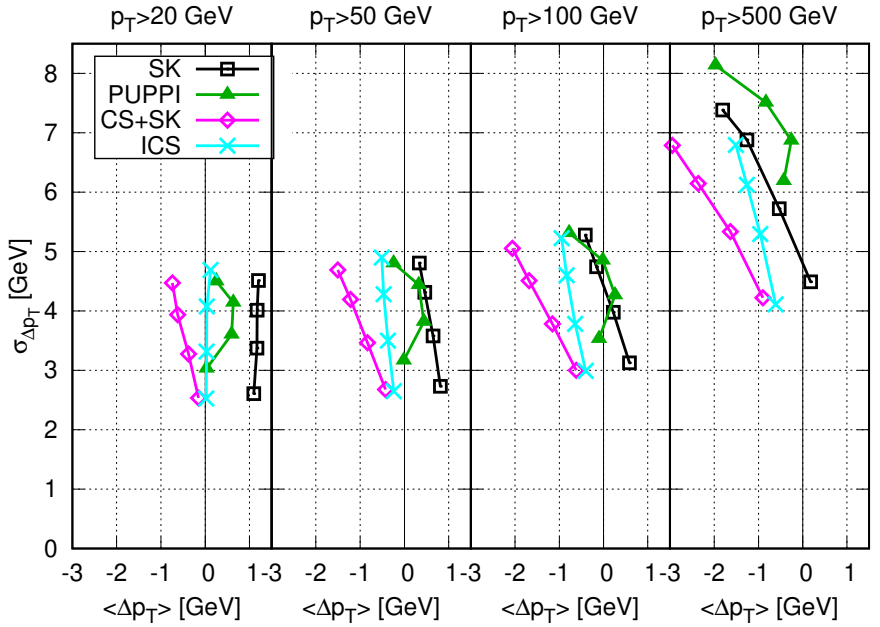


Figure 9. Comparison of the jet p_{T} resolution as a function of jet p_{T} bias for four methods: SoftKiller, PUPPI, CS+SoftKiller and ICS. Dijet events are used with different jet p_{T} cut in each panel. Each curve corresponds to a different method and the 4 points on each curve correspond to $N_{\mathrm{PU}} = 30$, 60, 100 and 140 from bottom to top.

JHEP08(2019)175

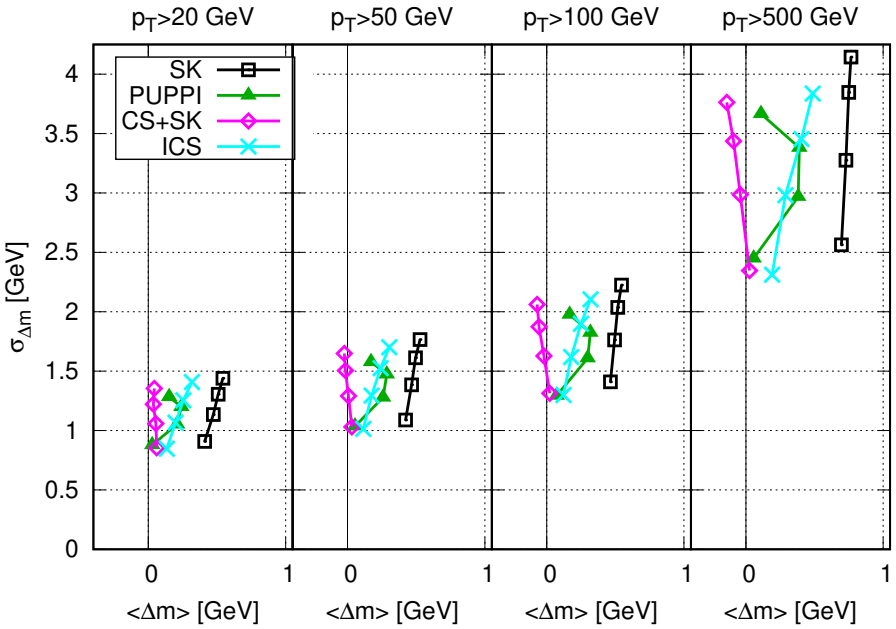


Figure 10. Comparison of the jet mass resolution as a function of jet mass bias for four methods: SoftKiller, PUPPI, CS+SoftKiller and ICS. Dijet events are used with different jet p_{T} cut in each panel. Each curve corresponds to a different method and the 4 points on each curve correspond to $N_{\mathrm{PU}} = 30$, 60, 100 and 140 from bottom to top.

servables or kinematic selections. For the majority of the N_{PU} and kinematic selections, the best performance is achieved by the ICS method. For all the N_{PU} and kinematic selections, PUPPI provides slightly worse resolution of p_{T} compared to ICS which is worsening with increasing p_{T}. On the contrary, for high N_{PU}, the performance of PUPPI is systematically better compared to ICS in terms of the reconstruction of mass. In the low N_{PU} environment the performance of ICS and CS+SoftKiller is almost identical, but for a higher N_{PU} environement, CS+SoftKiller gives larger negative bias. On the contrary, SoftKiller alone gives systematically positive bias. This oversubtraction or undersubtraction of SoftKiller might be avoided by further optimizing the SoftKiller parameters to avoid cutting out a part of the signal or to allow for cutting out more of the background. The SoftKiller performance may also be improved by applying the *protected zeroing* which removes all neutral particles below certain p_{T} threshold if a given particle is not located close to a charged particle from a hard-scattering process as discussed in ref. [28]. We should emphasize that in this study, the only method which uses the information about charged particles in the subtraction is PUPPI. In general, using the information from charged particles is expected to improve the performance of the subtraction. Systematic study of the methods which use the information from charged particles within the context of constituent-subtraction-based algorithms goes beyond the scope of this paper and is planned for a separate study.

JHEP08(2019)175

5 Conclusions

We presented a new background mitigation method for jet kinematics, jet substructure observables, and event observables (such as missing transverse energy), called Iterative Constituent Subtraction. This method is applicable to both pileup effects in proton-proton collisions and underlying event contributions in heavy-ion collisions. Iterative Constituent Subtraction extends the Constituent Subtraction method from ref. [8] to an iterative event-wide subtraction algorithm with improved features and performance. The new algorithm has been tested using hadronic jets from $Z' \to t\bar{t}$, as well as light quark and gluon dijet processes, and it was compared to various other algorithms including Jet-by-jet CS, Event-wide CS, Area Subtraction, SoftKiller, and PUPPI. Iterative Constituent Subtraction has been shown to significantly improve the performance of pileup subtraction in proton-proton collisions in terms of bias and resolution of the jet kinematics and substructure observables compared to the Area Subtraction, Jet-by-jet CS, and Event-wide CS. The improvement was also observed with respect to other methods for a large number of pileup and kinematic configurations. The new method has potential to improve the background mitigation at both proton-proton and heavy-ion colliders such as the LHC and RHIC.

A General discussion on the CS parameters

As discussed in section 2, the CS procedure has three free parameters to control the subtraction: $\Delta R^{\max}$, α, and A_g. These parameters have weak impact in the Jet-by-jet CS, and the recommended values in that configuration are $\Delta R^{\max} = \infty$, $\alpha = 0$, and $A_g \leq 0.01$. On the contrary, the Event-wide CS is much more sensitive to the $\Delta R^{\max}$ parameter and also the α parameter can have non-negligible effect. This section serves as a general guidance how to set the three parameters for the Event-wide CS. However, the conclusions found in our simulation may not be perfect for a specific detector environment. Therefore the experiments are encouraged to optimize the parameters in their own environment. The recommended starting point of tests for Event-wide CS is: $\Delta R^{\max} = 0.25$, $\alpha = 1$, $A_g = 0.0025$ for anti-k_t $R = 0.4$ jets and $\Delta R^{\max} = 0.7$, $\alpha = 1$, $A_g = 0.0025$ for anti-k_t $R = 1.0$ jets. We provide justification for these recommendations and a basic analysis of the sensitivity to the choice of parameters in the following sub-sections. The performance studies presented here use the setup described in section 4.3.

A.1 Maximal distance between ghost-particle pairs, $\Delta R^{\max}$

The parameter $\Delta R^{\max}$ controls the maximal allowed distance between ghost-particle pairs. Only ghost-particle pairs which have $\Delta R < \Delta R^{\max}$ are combined in the algorithm. By setting a finite $\Delta R^{\max}$, one can avoid combining ghost-particle pairs which are too far from each other. However, in this case, certain amount of estimated pileup can remain in the event, while by setting $\Delta R^{\max} = \infty$ it is ensured that all the estimated pileup represented by ghosts is subtracted from particles, although not necessarily at a naturally small distance between ghosts and particles.

JHEP08(2019)175

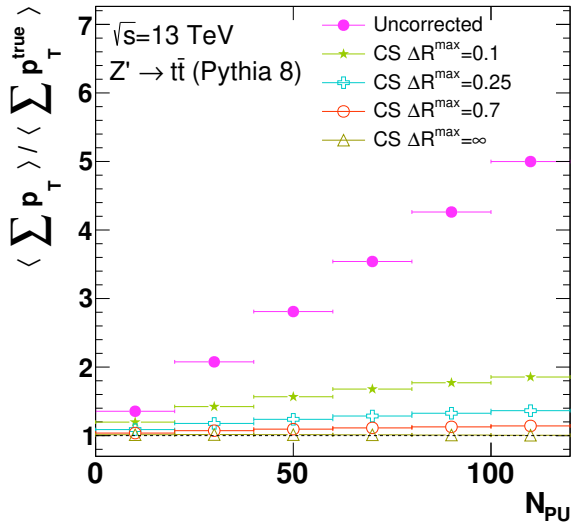


Figure 11. Ratio between the $\langle\sum p_{\mathrm{T}}\rangle$ for various ΔR^{max} parameters used in the Event-wide CS and $\langle\sum p_{\mathrm{T}}^{\mathrm{true}}\rangle$ as a function of N_{PU}.

We investigated the amount of remaining pileup for various ΔR^{max} values by evaluating the scalar p_{T} sum from all particles in the event. The average scalar p_{T} sum for true events (events without pileup) for the used physics process is $\langle\sum p_{\mathrm{T}}^{\mathrm{true}}\rangle \approx 1\,\mathrm{TeV}$. The average scalar p_{T} sum for the same events with pileup after correction, $\langle\sum p_{\mathrm{T}}\rangle$, varies depending on the ΔR^{max} parameter as shown in figure 11. For $\Delta R^{\mathrm{max}} = \infty$, the $\langle\sum p_{\mathrm{T}}\rangle \approx \langle\sum p_{\mathrm{T}}^{\mathrm{true}}\rangle$ for the whole N_{PU} range, which means that on average, the expected pileup deposition is entirely subtracted. This is expected, but this evaluation represents an important cross-check to see that the used background estimation is not biased. For finite ΔR^{max}, certain amount of pileup remains in the events. For example, there is still a remaining average background p_{T} of ∼80 MeV per unit area per one pileup event when using $\Delta R^{\mathrm{max}} = 0.25$.

The justification for using small values of ΔR^{max} comes from the actual performance on jets, see figures 12 and 13. When using $\Delta R^{\mathrm{max}} = \infty$, the jets are largely overcorrected. The reason for this are the fluctuations of pileup in the $y - \phi$ space. In case of no fluctuations, the ghosts subtract the p_{T} added by pileup rather accurately. However, the presence of fluctuations can cause that ghosts are more often matched to a hard-scatter particle, which may be distant, which leads to the overcorrection.

The optimal ΔR^{max} value may depend on the jet definition and the detector granularity. We found that $\Delta R^{\mathrm{max}} = 0.25$ leads to optimal performance for both anti-k_t $R = 0.4$ and anti-k_t $R = 1.0$ keeping the biases low and maximizing the resolution for most observables. Using $\Delta R^{\mathrm{max}} = 0.7$ can lead to lower biases for anti-k_t $R = 1.0$ jets in expense of worse resolution. In any case, it is recommended to apply an additional correction to remove the remaining pileup which can be addressed by the Iterative Constituent Subtraction method described in section 3. Other possibility is to use a different method. For example, SoftKiller (grid parameter of 0.6) applied after the Event-wide CS with $\Delta R^{\mathrm{max}} = 0.25$ was found to be one of the best methods in studies from the ATLAS Collaboration [40].

JHEP08(2019)175

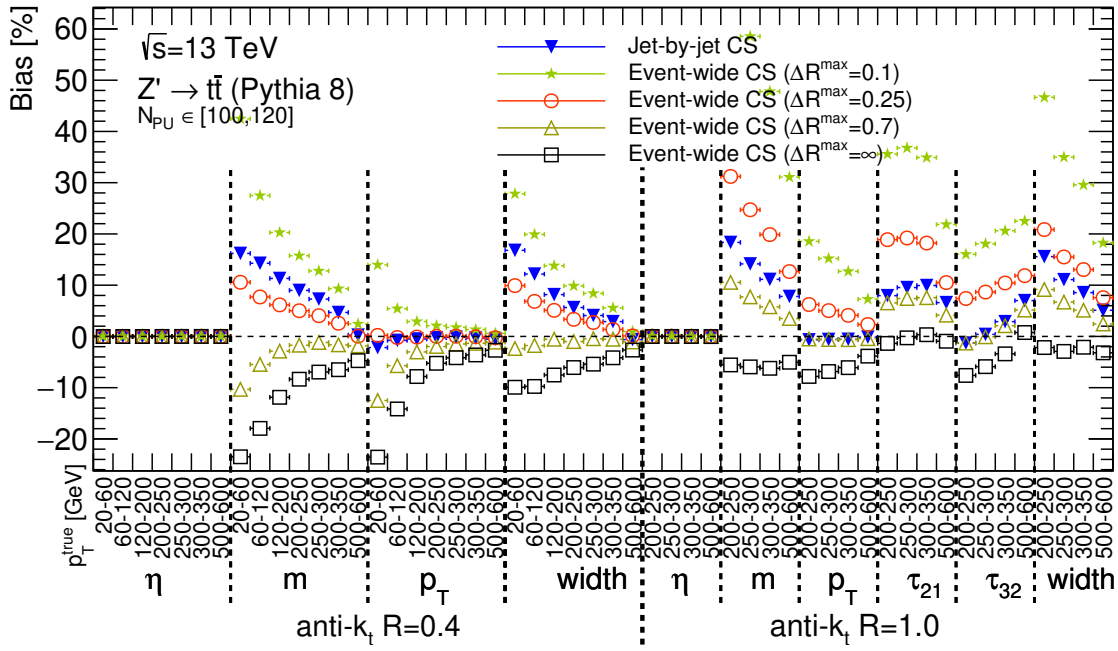


Figure 12. Dependence of bias on the $\Delta R^{\max}$ parameter for the Event-wide CS. The other CS parameters are $\alpha = 0$ and $A_g = 0.0025$.

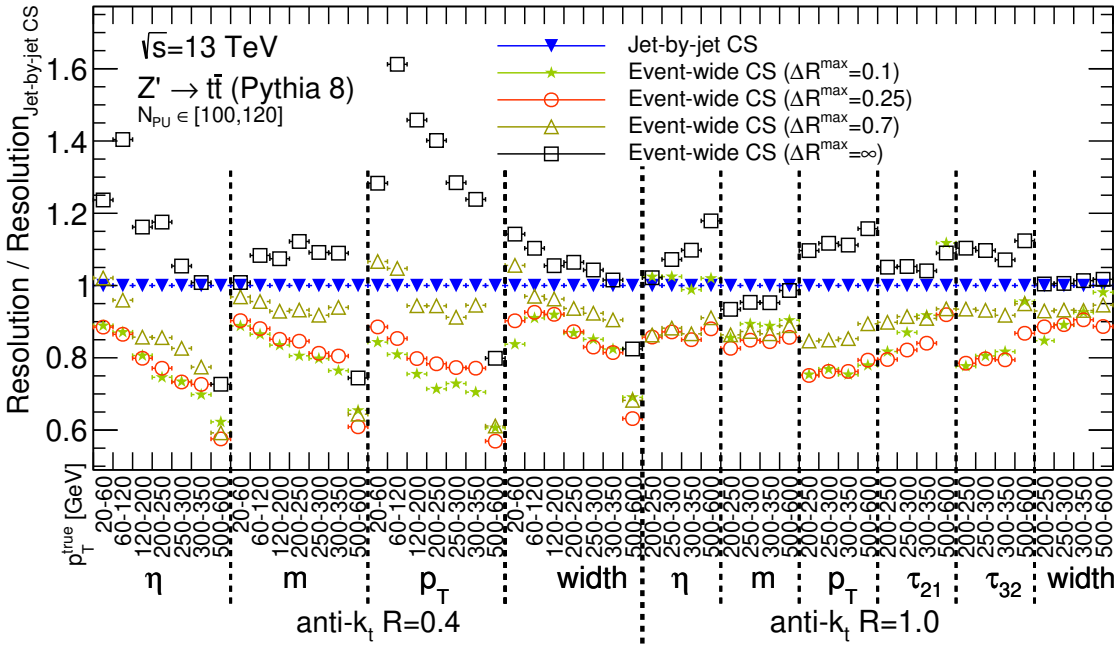


Figure 13. Dependence of resolution on the $\Delta R^{\max}$ parameter for the Event-wide CS. The other CS parameters are $\alpha = 0$ and $A_g = 0.0025$.

JHEP08(2019)175

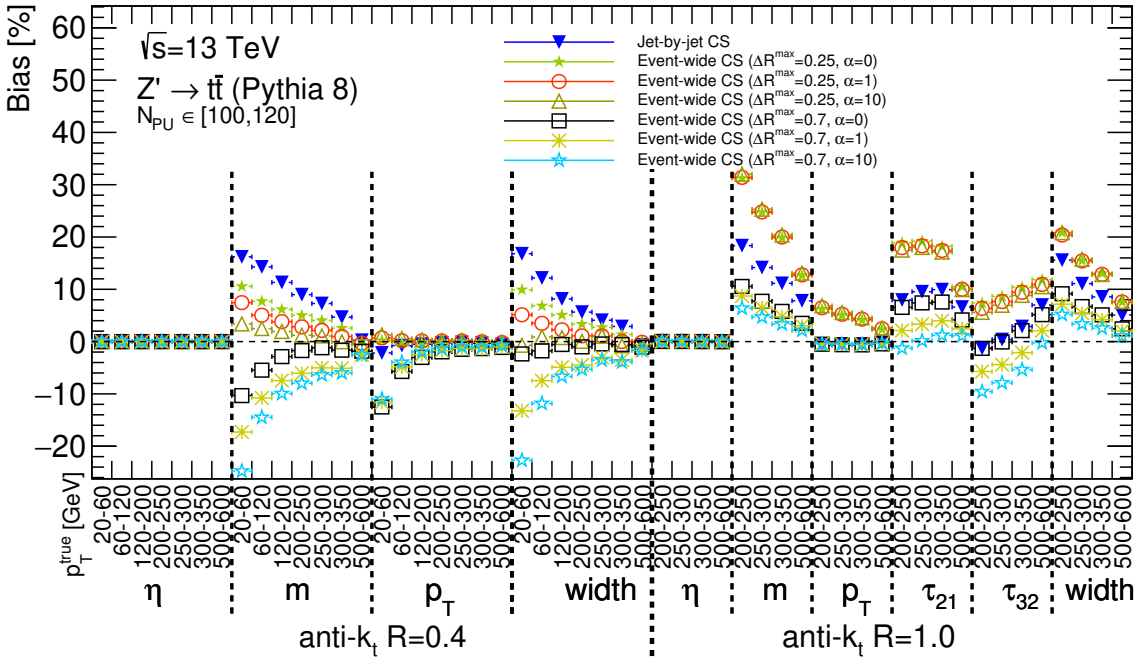


Figure 14. Dependence of bias on the α parameter for the Event-wide CS.

A.2 α parameter

By using $\alpha > 0$, one can prioritize ghost-particle pairs with lower particle p_{T} during the subtraction procedure. This may have positive effect on the performance since it is expected that the pileup particles have lower p_{T} than the hard-scatter particles. The optimal value of the α parameter depends on many factors: granularity of the detector, possible p_{T} cuts applied to all particles, jet definition, ΔR^{max} parameter and also on the observable in question.

The dependence of the Event-wide CS performance on α is shown in figures 14 and 15. In general, the smaller the ΔR^{max}, the smaller is the effect of the α parameter. This is expected since with smaller ΔR^{max}, each ghost has smaller freedom to move to match with a particle during the CS procedure. The choice of α parameter has almost no effect for the anti-k_t $R = 1.0$ jets when using small values of ΔR^{max} (e.g. $\Delta R^{\mathrm{max}} = 0.25$). In this case, each ghost is active over an area which is much smaller than the area of the jet, that is locally, and the choice of α has no real impact.

We found that the choice $\alpha = 1$ for anti-k_t $R = 0.4$ and $\Delta R^{\mathrm{max}} = 0.25$ keeps the biases low while the resolution is maximized. For anti-k_t $R = 1.0$ jets, $\Delta R^{\mathrm{max}} = 0.7$ and $\alpha = 1$ is preferable.

A.3 Ghost area A_g

The smaller the value of A_g, the more densely are the ghosts distributed in the $y - \phi$ space which leads to better performance of the CS procedure. On the other hand, the smaller the value of A_g, the higher is the number of ghosts which implies a longer computational time. Therefore in practice, a compromise needs to be done. The optimal value of A_g

JHEP08(2019)175

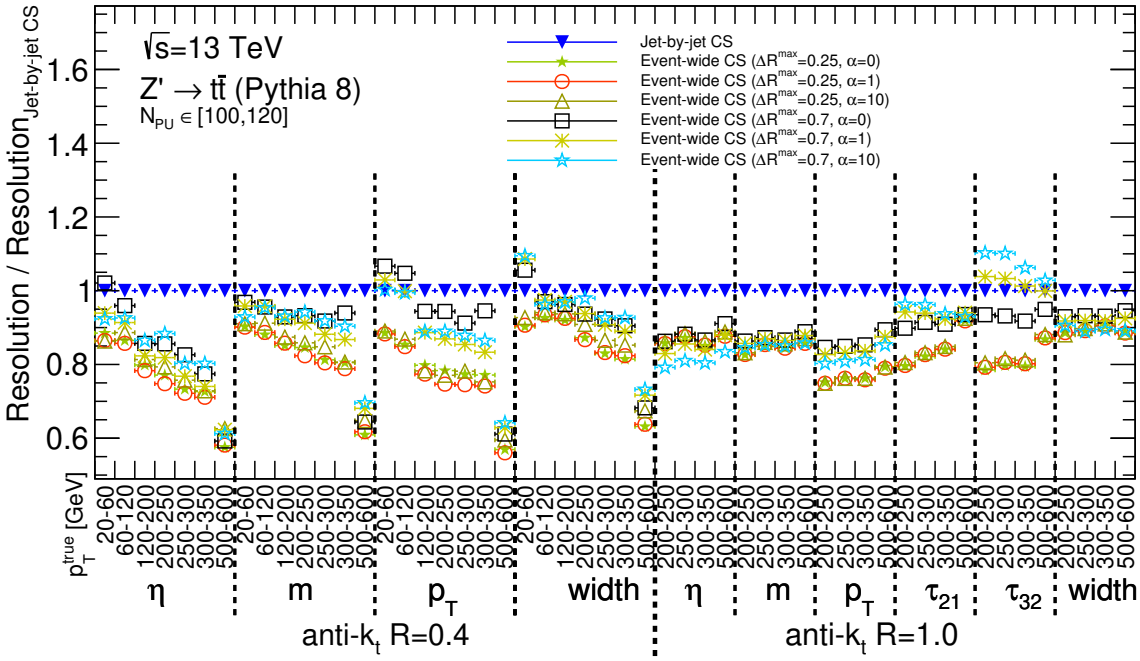


Figure 15. Dependence of resolution on the α parameter for the Event-wide CS.

may depend on the jet definition, granularity of the detector, and the $\Delta R^{\max}$ parameter in the CS procedure. We found that $A_g = 0.0025$ gives the best performance and using smaller A_g than 0.0025, does not lead to any significant improvement, it just brings a longer computational time. By using $A_g = 0.01$, we observe on average 4-times faster correction with subtle worsening of performance (relatively up to ∼2% worse resolution for some observables).

B Choice of parameters for the ICS method

We investigated the optimization of configuration of the ICS algorithm by varying: number of iterations, CS parameters ($\Delta R^{\max}$, α, A_g) for each iteration, and the usage of ghost removal option. The performance studies presented here use the setup described in section 4.3. We found that the best performance for the majority of studied observables and jet p_{T} ranges can be achieved with two iterations using ghost removal and parameters $\alpha = 1$, $A_g = 0.0025$, while the $\Delta R^{\max}$ parameter is different for each iteration and also depends on the jet definition. For anti-k_t $R = 0.4$ jets, parameters $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.1$ are found to be optimal for the first and second iteration, respectively. For anti-k_t $R = 1.0$ jets, the parameters $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.35$ are found to be optimal. We would like to emphasize, that these conclusions may not be perfect for a specific detector environment and/or for the case of heavy-ion collisions. Therefore, the experiments are encouraged to optimize the ICS configuration in their own environment.

In the following we provide a justification for the choice of the optimal ICS configuration based on performance studies. First we discuss ICS with two iterations and then we focus on the performance when using three or more iterations. The discussion of the

JHEP08(2019)175

choice of CS parameters $\Delta R^{\max}$, α, A_g, which is provided in appendix A, applies also for the ICS method. We use ghost area $A_g = 0.0025$ and $\alpha = 1$ in the following, since we have not observed any significant improvement by using smaller A_g or different α parameter. Therefore, we discuss just the parameter $\Delta R^{\max}$, the option of ghost removal, and the number of iterations in the following.

B.1 ICS with two iterations

In general, it is expected that the ghost removal should help for cases when $\Delta R_1^{\max} \geq \Delta R_2^{\max}$, since then the ghost removal can guarantee that all the expected background deposition is subtracted (in case of not using ghost removal, ghosts, which were not fully subtracted in the first iteration, have no chance to be subtracted in the second iteration). For cases when $\Delta R_1^{\max} < \Delta R_2^{\max}$, all the expected background can be subtracted independently of the usage of the ghost removal option. The impact of the ghost removal may depend on properties of the background. In case when ρ is correlated in close-by regions in the $y - \phi$ space, then it may be beneficial to use the ghost removal. For example, when one ghost is not fully subtracted in the first iteration, it means that the real ρ in the vicinity of this ghost is smaller than it was originally estimated. Therefore when the ρ is correlated in close-by regions, then the ρ within larger area around this ghost should be also smaller than the estimated ρ so it may be helpful to remove such ghosts. On the other hand, when the ρ in close-by regions is anti-correlated, then it makes sense not to use the ghost removal.

We investigated the performance of ICS with two iterations for reasonable values of $\Delta R^{\max}$ parameter taking into account the findings in appendix A.1. For this, we used 50 configurations which were obtained by using $\Delta R^{\max}$ parameter for the first iteration from the set $\{0.1, 0.15, 0.2, 0.35, 0.7\}$ and for the second iteration from the set $\{0.1, 0.15, 0.2, 0.35, \infty\}$, and using or not using ghost removal. The bias and resolution for these 50 configurations vary significantly for the studied jet observables and jet p_{T} ranges. For each jet observable and jet p_{T} range, there is configuration which have maximal resolution at the expense of having higher bias compared to other configurations. To define the optimal configuration, we rather focus on minimizing the bias. Among the configurations which have p_{T} bias less than 3% and bias for other observables less than 8%, we define the optimal configuration for each jet definition as the one which has maximal resolution for the highest number of jet observables and jet p_{T} bins. Based on this definition, we found that using the ghost removal option together with parameters $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.1$ for anti-k_t $R = 0.4$ jets, and parameters $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.35$ for anti-k_t $R = 1.0$ jets provides the optimal configuration. We show the jet performance for a representative set of ICS configurations in figures 16 and 17.

First we discuss the performance on anti-k_t $R = 0.4$ jets. The two shown configurations with $\Delta R_1^{\max} = 0.1$ and $\Delta R_2^{\max} = 0.1$ have the best resolution, but high bias for jet mass and jet width, especially for low-p_{T} ranges. We observe that the usage of the ghost removal option reduces the bias while the resolution stays similar for these $\Delta R^{\max}$ values. The bias for jet mass and jet width can be further reduced by using ghost removal with $\Delta R_1^{\max} = 0.2$

JHEP08(2019)175

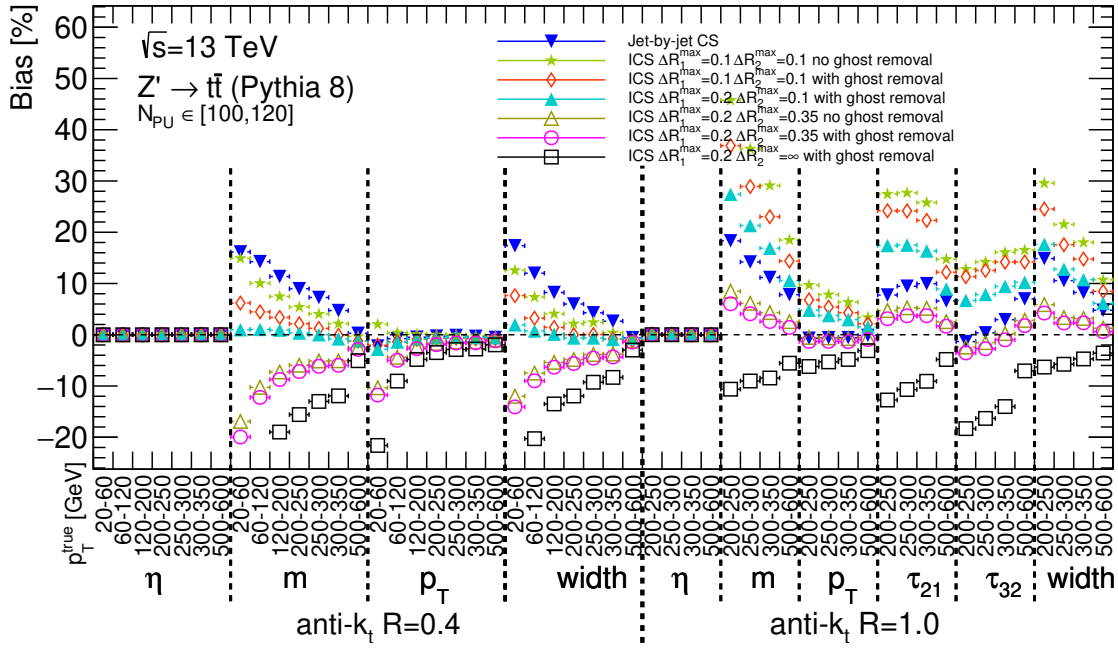


Figure 16. Bias for a representative set of configurations for the ICS method with two iterations. The $\Delta R^{\max}$ parameter for the individual iterations is varied and the ghost removal option is used or not. The other CS parameters are set to $\alpha = 1$ and $A_g = 0.0025$ for each iteration.

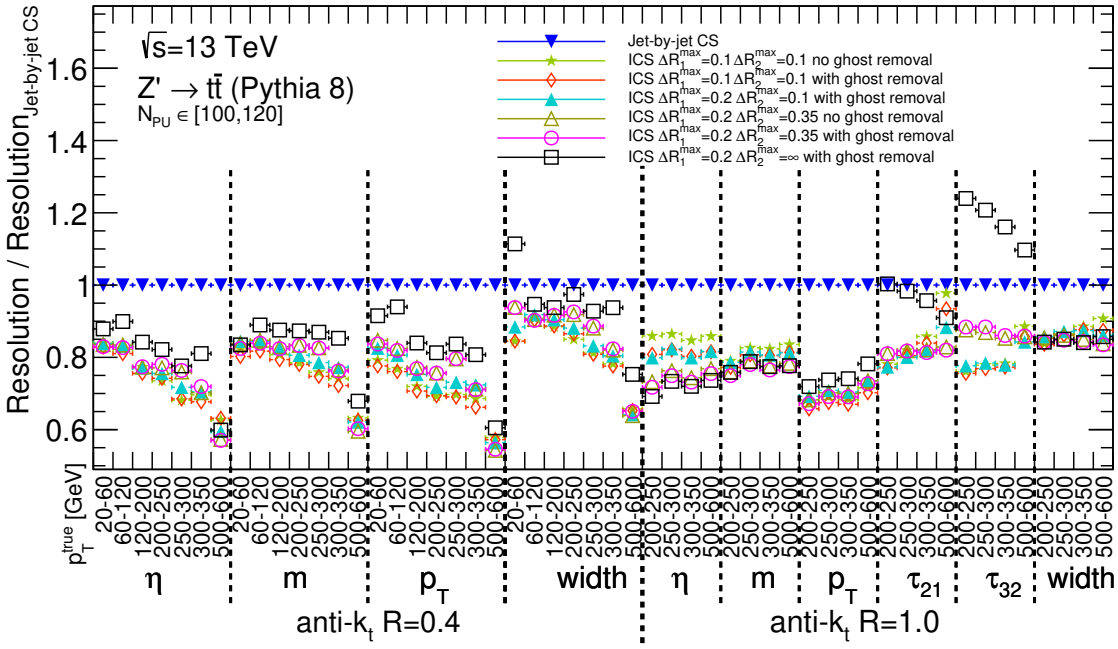


Figure 17. Resolution for a representative set of configurations for the ICS method with two iterations. The $\Delta R^{\max}$ parameter for the individual iterations is varied and the ghost removal option is used or not. The other CS parameters are set to $\alpha = 1$ and $A_g = 0.0025$ for each iteration.

JHEP08(2019)175

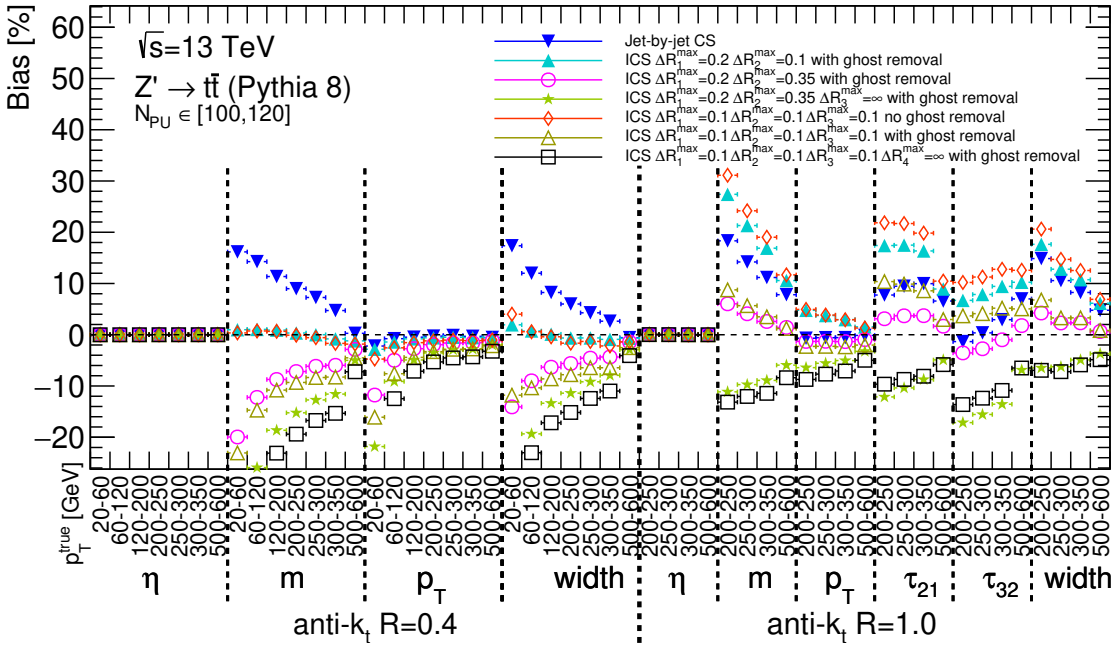


Figure 18. Bias for a representative set of configurations for the ICS method with three or more iterations. The $\Delta R^{\max}$ parameter for the individual iterations is varied and the ghost removal option is used or not. The other CS parameters are set to $\alpha = 1$ and $A_g = 0.0025$ for each iteration.

and $\Delta R_2^{\max} = 0.1$ which is the configuration that we identified as the optimal one. By using higher $\Delta R_2^{\max}$, the bias and resolution for all jet observables get worse.

To get the optimal performance for anti-k_t $R = 1.0$ jets higher $\Delta R^{\max}$ values are preferred than for anti-k_t $R = 0.4$ jets, which is consistent with findings in appendix A.1. The configuration with $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = 0.35$ is found to be optimal. The ghost removal then further slightly improves the performance. The effect of ghost removal is much smaller for these $\Delta R^{\max}$ parameters compared to values $\Delta R_1^{\max} = 0.1$ and $\Delta R_2^{\max} = 0.1$, which is expected since in the first iteration larger fraction of ghost p_{T} is subtracted so less scalar p_{T} is redistributed in the second iteration. The configuration $\Delta R_1^{\max} = 0.2$ and $\Delta R_2^{\max} = \infty$ ensures that all expected background is subtracted, however it still seems that at wrong places, since the performance is worse for this configuration. At the same time the performance is still better compared to the performance of simple Event-wide CS with $\Delta R^{\max} = \infty$.

B.2 ICS with three or more iterations

We tried several configurations with three or more iterations. Based on the findings in the previous subsection, we used $\Delta R_1^{\max}$ and $\Delta R_2^{\max}$ values which exhibited only small overcorrection of jet p_{T} as a starting point. The jet performance for a representative set of configurations is shown in figures 18 and 19. We did not find any significant improvement when using three or more iterations. However, some jet observables can have by ∼10% better resolution at the expense of slightly worse bias when using more iterations. This

JHEP08(2019)175

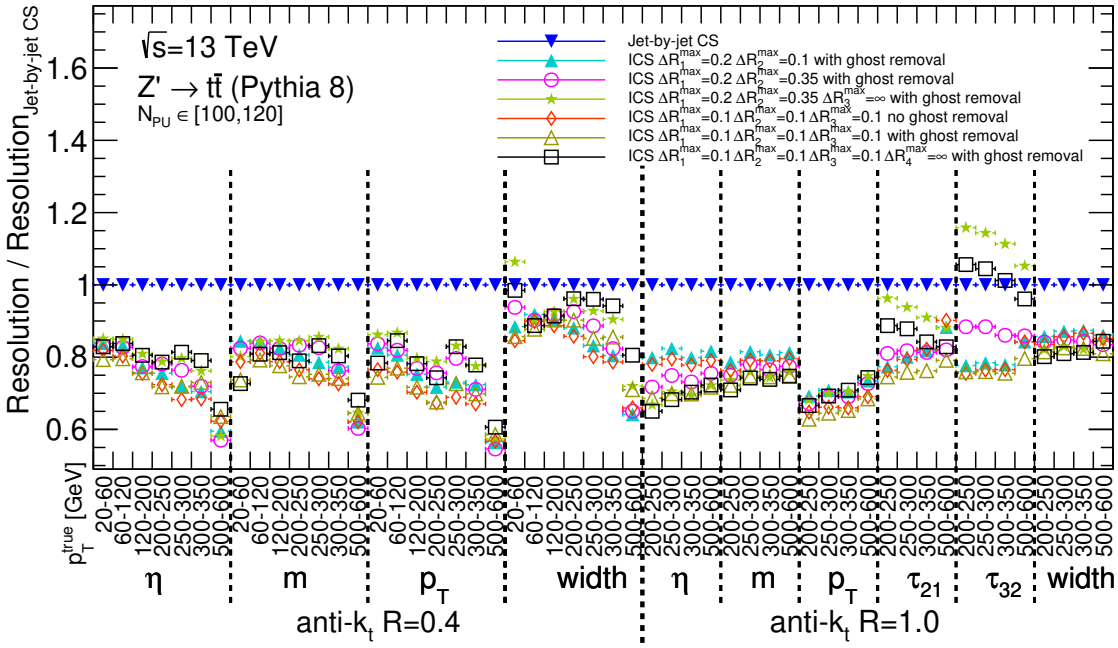


Figure 19. Resolution for a representative set of configurations for the ICS method with three or more iterations. The $\Delta R^{\max}$ parameter for the individual iterations is varied and the ghost removal option is used or not. The other CS parameters are set to $\alpha = 1$ and $A_g = 0.0025$ for each iteration.

holds especially for the configuration with three iterations using $\Delta R_1^{\max} = 0.1$, $\Delta R_2^{\max} = 0.1$, and $\Delta R_3^{\max} = 0.1$. Similarly as in the previous section, we observe that forcing to subtract all expected background by setting $\Delta R_3^{\max} = \infty$ is worse than keeping $\Delta R_3^{\max}$ finite (see e.g. configuration with $\Delta R_1^{\max} = 0.2$, $\Delta R_2^{\max} = 0.35$, and $\Delta R_3^{\max} = \infty$).

In this appendix, we identified and discussed an optimal configuration for which ICS preforms better than other background mitigation methods (see section 4) and provided a basic insight into the mechanisms that drives the performance. At the same time, the presented optimal configuration should not be viewed as the best configuration for all experiments and all backgrounds. The presented configuration should be viewed only as an advice which can be used as a starting point for further optimization within realistic environment of experiments.

C Treatment of massive particles

In the case of massless particles, there are three degrees of freedom for each particle 4-momentum. The CS procedure corrects p_{T} of each particle, while the rapidity and azimuth are kept unchanged. In that way, it is ensured that the anti-k_t jet clustering algorithm clusters the same particles.[5] Since $y = \eta$ for massless particles, also the momentum direc-

[5]Up to small back-reaction biases coming from non-perfect correction which will be neglected in the following.

JHEP08(2019)175

tion is kept unchanged which ensures good reconstruction of some jet observables such as the jet mass.

In the case of massive particles, there are four degrees of freedom for each particle 4-momentum. The CS procedure corrects the p_{T} of each particle to compensate for the presence of background particles on average. Then the particle azimuth should be kept unchanged since there is no reason why the presence of background would change the azimuth. However, it is non-trivial to decide how to treat the remaining two degrees of freedom for the particle 4-momentum. One straightforward option is to keep the mass of each particle unchanged. For the remaining degree of freedom, two obvious options are available: keep rapidity or pseudo-rapidity unchanged. Due to the fact that the anti-k_t jet clustering algorithm uses difference of rapidities in the definition of the distance measure, there are the following caveats for the two options:

1. Keeping mass and rapidity unchanged — in this case, it is ensured that the difference of rapidities in the anti-k_t algorithm of each particle pair is not modified after the correction, so the particle content of the final jets can be similar to the particle content of true jets. However, in the limit when the corrected particle $p_{\mathrm{T}} \to 0$, the z-component of momentum $p_z \to m \sinh y$ and pseudo-rapidity $\eta \to \pm\infty$, which means that after such correction, the central anti-k_t jets may contain particles with non-negligible momentum pointing to totally different direction than the high p_{T} particles in that jet. This causes a very large bias for some jet observables such as jet mass.

2. Keeping mass and pseudo-rapidity unchanged — in this case, the difference of rapidities in the anti-k_t algorithm of each particle pair can be modified which may lead to jets with different particle content compared to the true jets. Especially, in the limit when the corrected particle $p_{\mathrm{T}} \to 0$, the rapidity $y \to 0$ and the energy $E \to m$. Therefore after the correction, jets near $\eta = 0$ typically contain a lot of constituents with non-zero mass but very small energy.

To summarize these two options we can say that although the particle mass is usually negligible compared to the particle energy, we see that it cannot stay unchanged due to the fact that the anti-k_t jet clustering algorithm uses rapidity difference in its distance measure. Hence, it is important to modify the mass of the particles. There are several other options how to treat the two remaining degrees of freedom:

3. Set the particle masses to zero. Keep the rapidity unchanged.

4. Set the particle masses to zero. Keep the pseudo-rapidity unchanged.

5. Do correction of variable $m_\delta = \sqrt{p_{\mathrm{T}}^2 + m^2} - p_{\mathrm{T}}$ in the same way as it is described in the original CS procedure [8] which was based on the approach in ref. [7]. Keep the rapidity unchanged.

6. Do the same correction of m_δ as in the previous option, just keep the pseudo-rapidity instead of rapidity unchanged.

JHEP08(2019)175

7. Keep the rapidity and pseudo-rapidity unchanged. This option is equivalent with scaling the original 4-momentum by a factor corresponding to the ratio between the corrected and original p_{T} of the particle.

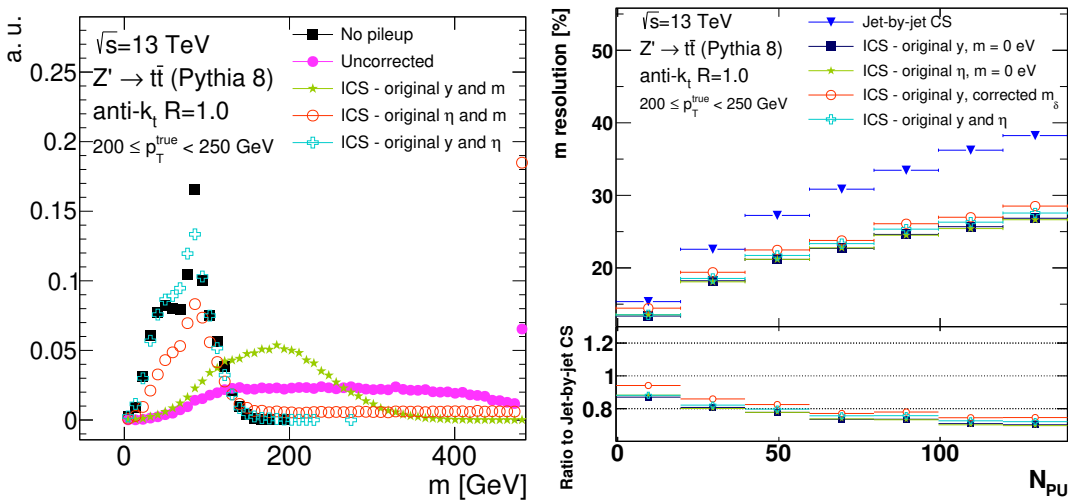
We note that by setting the particle mass to zero (options 3 and 4), a bias is introduced for $N_{\mathrm{PU}} = 0$ by definition (since all the hard-scatter particles are modified to be massless). There is no such bias for options $5-7$. The options 5 and 6, which use the m_δ correction, can still suffer from the same effects as described above the for options 1 and 2, respectively, thought with smaller impact on the performance. The options 3, 4, and 7 are not affected by the effects described above for options 1 and 2.

We investigated the performance of the individual options using the ICS method in the same conditions as described in section 4.3 but without the detector simulation. For these studies, we used ICS from `FastJet Contrib` version 1.042, where all the options for the treatment of massive particles are implemented. We found that the jet performance for options 1 and 2 is much worse than for options $3-7$, especially for the jet mass. The performance for jet η, jet p_{T}, jet width and subjettiness ratios is very similar among the options $3-7$. From the tested jet observables, only the jet mass has large dependence on the chosen strategy for correcting the particle 4-momenta, therefore we discuss the performance for the jet mass in the following.

First, the importance of modifying the particle mass is demonstrated in figure 20a showing the jet mass distribution for anti-k_t $R = 1.0$ jets with true $p_{\mathrm{T}} = 200 - 250\,\mathrm{GeV}$ (containing mainly decay products from the W boson). The true distribution is peaked around the W boson mass. With pileup, the original distribution is broadened and shifted to higher values. After ICS correction with keeping the original mass and rapidity of particles (option 1), the jet mass distribution is partially corrected, but still it is significantly shifted away from the true distribution. When using ICS correction with keeping the original mass and pseudo-rapidity (option 2), the corrected distribution agrees with the true distribution a bit better, but still it has a significant tail from jets biased towards high jet mass. In contrast, when using one of the options without keeping the original mass, e.g. option 7 (keeping original rapidity and pseudo-rapidity), the corrected mass distribution agrees much better with the true distribution.

Next, we investigated the bias and resolution of the jet mass systematically for options $3-7$. We found that option 6 has bias above 20% and much worse resolution than the other four options. The option 5 has bias up to 10% and resolution comparable with the options 3, 4, and 7. These remaining options have bias below 2%. The jet mass resolution for all these options using ICS as a function of N_{PU} is shown in figure 20b compared to jet-by-jet CS (which uses option 3). As demonstrated in this figure, we found that setting the particle mass to zero (options 3 and 4) performs the best for the anti-k_t $R = 1.0$ jets with true $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, while option 7 has only slightly worse resolution and option 5 has worse resolution by $5-10\%$. We found that the bias introduced by setting the particle mass to zero (options 3 and 4) gets negligible for $N_{\mathrm{PU}} > 5$. We also observed that this jet mass bias is larger for low p_{T} anti-k_t $R = 0.4$ jets where the option 7 performs slightly better in general compared to options 3 and 4.

JHEP08(2019)175



(a) (b)

Figure 20. Jet mass distribution (left) for the true jets, jets with pileup, and ICS corrected jets for three different approaches for the treatment of massive inputs: two approaches in which the particle mass is unchanged and one approach in which the rapidity and pseudo-rapidity is kept unchanged. Jet mass resolution (right) for six correction methods. The last bin is the overflow bin.

We conclude from our studies, that keeping the original rapidity and pseudo-rapidity (option 7) leads to the optimal performance for the two studied jet definitions over the given phase space. However, the experiments using massive inputs are encouraged to check our findings for the exact setup of the analysis they intend to perform.

There is one other aspect of the background subtraction with massive particles. As mentioned in section 2, a user-defined parameter $\eta^{\max}$ is used, which ensures that only particles with $|\eta| < \eta^{\max}$ are corrected by constructing the ghosts up to the same $|\eta|$ limit. However, the available software for ρ estimation in `FastJet` uses particle rapidities and not pseudo-rapidities. Therefore, in the case of massive particles, it is important that the user uses only particles with $|\eta| < \eta^{\max}$ to estimate the ρ and also to derive the rapidity dependence used for background rescaling.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 797520. The work of MS was supported by Grant Agency of the Czech Republic under Grant 18-12859Y and by Charles University grants UNCE/SCI/013 and Progres Q47. DWM is supported by the National Science Foundation under Grant No. PHY-1454815.

JHEP08(2019)175

References

[1] ATLAS luminosity public results from run 2, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (2018).

[2] CMS luminosity public results from run 2, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults (2018).

[3] ATLAS and CMS collaborations, *Report on the Physics at the HL-LHC and Perspectives for the HE-LHC*, in *HL/HE-LHC Physics Workshop: final jamboree*, CERN, Geneva Switzerland (2019).

[4] CMS collaboration, *Observation and studies of jet quenching in PbPb collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV*, *Phys. Rev.* **C 84** (2011) 024906 [arXiv:1102.1957] [INSPIRE].

[5] M. Cacciari and G.P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett.* **B 659** (2008) 119 [arXiv:0707.1378] [INSPIRE].

[6] M. Cacciari, G.P. Salam and G. Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005 [arXiv:0802.1188] [INSPIRE].

[7] G. Soyez, G.P. Salam, J. Kim, S. Dutta and M. Cacciari, *Pileup subtraction for jet shapes*, *Phys. Rev. Lett.* **110** (2013) 162001 [arXiv:1211.2811] [INSPIRE].

[8] P. Berta, M. Spousta, D.W. Miller and R. Leitner, *Particle-level pileup subtraction for jets and jet shapes*, *JHEP* **06** (2014) 092 [arXiv:1403.3108] [INSPIRE].

[9] M. Cacciari, G.P. Salam and G. Soyez, *SoftKiller, a particle-level pileup removal method*, *Eur. Phys. J.* **C 75** (2015) 59 [arXiv:1407.0408] [INSPIRE].

[10] D. Bertolini, P. Harris, M. Low and N. Tran, *Pileup Per Particle Identification*, *JHEP* **10** (2014) 059 [arXiv:1407.6013] [INSPIRE].

[11] D. Krohn, M.D. Schwartz, M. Low and L.-T. Wang, *Jet Cleansing: Pileup Removal at High Luminosity*, *Phys. Rev.* **D 90** (2014) 065020 [arXiv:1309.4777] [INSPIRE].

[12] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [arXiv:0802.2470] [INSPIRE].

[13] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084 [arXiv:0912.1342] [INSPIRE].

[14] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, *Phys. Rev.* **D 81** (2010) 094023 [arXiv:0912.0033] [INSPIRE].

[15] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146 [arXiv:1402.2657] [INSPIRE].

[16] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [arXiv:1307.0007] [INSPIRE].

[17] F.A. Dreyer, L. Necib, G. Soyez and J. Thaler, *Recursive Soft Drop*, *JHEP* **06** (2018) 093 [arXiv:1804.03657] [INSPIRE].

[18] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323] [INSPIRE].

JHEP08(2019)175

[19] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Monte Carlo generators for HERA physics. Proceedings of Workshop*, Hamburg Germany (1998), pg. 270.

[20] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B 406** (1993) 187 [INSPIRE].

[21] S.D. Ellis and D.E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D 48** (1993) 3160 [hep-ph/9305266] [INSPIRE].

[22] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189] [INSPIRE].

[23] P.T. Komiske, E.M. Metodiev, B. Nachman and M.D. Schwartz, *Pileup Mitigation with Machine Learning (PUMML)*, *JHEP* **12** (2017) 051 [arXiv:1707.08600] [INSPIRE].

[24] J. Arjona Martínez, O. Cerri, M. Pierini, M. Spiropulu and J.-R. Vlimant, *Pileup mitigation at the Large Hadron Collider with Graph Neural Networks*, *Eur. Phys. J. Plus* **134** (2019) 333 [arXiv:1810.07988].

[25] S. Carrazza and F.A. Dreyer, *Jet grooming through reinforcement learning*, *Phys. Rev.* **D 100** (2019) 014014 [arXiv:1903.09644] [INSPIRE].

[26] M. Cacciari, G.P. Salam and G. Soyez, *Use of charged-track information to subtract neutral pileup*, *Phys. Rev.* **D 92** (2015) 014003 [arXiv:1404.7353] [INSPIRE].

[27] Y. Mehtar-Tani, A. Soto-Ontoso and M. Verweij, *A background estimator for jet studies in p+p and A+A collisions*, arXiv:1904.12815 [INSPIRE].

[28] G. Soyez, *Pileup mitigation at the LHC: a theorist's view*, habilitation, IPhT, Saclay France (2018).

[29] G. Bellettini, R. Bertani, C. Bradaschia, R. Del Fabbro, A. Scribano and G. Terreni, *Hadron calorimeter towers with a high space resolution*, *Nucl. Instrum. Meth.* **204** (1982) 73 [INSPIRE].

[30] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002 (2008).

[31] CMS collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, 2017 *JINST* **12** P10003 [arXiv:1706.04965] [INSPIRE].

[32] ALICE collaboration, *Medium modification of the shape of small-radius jets in central Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV*, *JHEP* **10** (2018) 139 [arXiv:1807.06854] [INSPIRE].

[33] ALICE collaboration, *First measurement of jet mass in Pb-Pb and p-Pb collisions at the LHC*, *Phys. Lett.* **B 776** (2018) 249 [arXiv:1702.00804] [INSPIRE].

[34] CMS collaboration, *Measurement of the groomed jet mass in PbPb and pp collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV*, *JHEP* **10** (2018) 161 [arXiv:1805.05145] [INSPIRE].

[35] CMS collaboration, *Measurement of the Splitting Function in pp and Pb-Pb Collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV*, *Phys. Rev. Lett.* **120** (2018) 142302 [arXiv:1708.09429] [INSPIRE].

[36] CMS collaboration, *Updates to Constituent Subtraction in Heavy Ions at CMS*, CERN-CMS-DP-2018-024 (2018).

[37] CMS collaboration, *Pileup Removal Algorithms*, CMS-PAS-JME-14-001 (2014).

JHEP08(2019)175

[38] ATLAS collaboration, *Impact of Alternative Inputs and Grooming Methods on Large-R Jet Reconstruction in ATLAS*, ATL-PHYS-PUB-2017-020 (2017).

[39] ATLAS collaboration, *Impact of Pile-up on Jet Constituent Multiplicity in ATLAS*, ATL-PHYS-PUB-2018-011 (2018).

[40] ATLAS collaboration, *Constituent-level pile-up mitigation techniques in ATLAS*, ATLAS-CONF-2017-065 (2017).

[41] M. Boronat, J. Fuster, I. Garcia, P. Roloff, R. Simoniello and M. Vos, *Jet reconstruction at high-energy electron–positron colliders*, *Eur. Phys. J.* **C 78** (2018) 144 [arXiv:1607.05039] [INSPIRE].

[42] T. Golling et al., *Physics at a 100 TeV pp collider: beyond the Standard Model phenomena*, *CERN Yellow Rep.* **3** (2017) 441.

[43] STAR collaboration, *Measurements of the jet internal sub-structure and its relevance to parton shower evolution in p+p and Au+Au collisions at STAR*, PoS(HardProbes2018)090 [arXiv:1903.12115].

[44] J.K. Behr, D. Bortoletto, J.A. Frost, N.P. Hartland, C. Issever and J. Rojo, *Boosting Higgs pair production in the $b\bar{b}b\bar{b}$ final state with multivariate techniques*, *Eur. Phys. J.* **C 76** (2016) 386 [arXiv:1512.08928] [INSPIRE].

[45] G.P. Salam, L. Schunk and G. Soyez, *Dichroic subjettiness ratios to distinguish colour flows in boosted boson tagging*, *JHEP* **03** (2017) 022 [arXiv:1612.03917] [INSPIRE].

[46] A.J. Larkoski, F. Maltoni and M. Selvaggi, *Tracking down hyper-boosted top quarks*, *JHEP* **06** (2015) 032 [arXiv:1503.03347] [INSPIRE].

[47] A. Berlin, T. Lin, M. Low and L.-T. Wang, *Neutralinos in Vector Boson Fusion at High Energy Colliders*, *Phys. Rev.* **D 91** (2015) 115002 [arXiv:1502.05044] [INSPIRE].

[48] N. Craig, H.K. Lou, M. McCullough and A. Thalapillil, *The Higgs Portal Above Threshold*, *JHEP* **02** (2016) 127 [arXiv:1412.0258] [INSPIRE].

[49] C. Brust, P. Maksimovic, A. Sady, P. Saraswat, M.T. Walters and Y. Xin, *Identifying boosted new physics with non-isolated leptons*, *JHEP* **04** (2015) 079 [arXiv:1410.0362] [INSPIRE].

[50] M. Low and L.-T. Wang, *Neutralino dark matter at 14 TeV and 100 TeV*, *JHEP* **08** (2014) 161 [arXiv:1404.0682] [INSPIRE].

[51] F.A. Dreyer, G.P. Salam and G. Soyez, *The Lund Jet Plane*, *JHEP* **12** (2018) 064 [arXiv:1807.04758] [INSPIRE].

[52] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C 72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[53] FastJet Contrib, February 2019.

[54] M. Cacciari and G.P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B 641** (2006) 57 [hep-ph/0512210] [INSPIRE].

[55] OPAL collaboration, *QCD studies using a cone based jet finding algorithm for e^+e^- collisions at LEP*, *Z. Phys.* **C 63** (1994) 197 [INSPIRE].

[56] D0 collaboration, *Transverse energy distributions within jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Lett.* **B 357** (1995) 500 [INSPIRE].

JHEP08(2019)175

[57] H1 collaboration, *Measurement of internal jet structure in dijet production in deep inelastic scattering at HERA*, *Nucl. Phys.* **B 545** (1999) 3 [hep-ex/9901010] [INSPIRE].

[58] ZEUS collaboration, *Substructure dependence of jet cross sections at HERA and determination of α_s*, *Nucl. Phys.* **B 700** (2004) 3 [hep-ex/0405065] [INSPIRE].

[59] CDF collaboration, *Study of jet shapes in inclusive jet production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev.* **D 71** (2005) 112002 [hep-ex/0505013] [INSPIRE].

[60] ATLAS collaboration, *Study of Jet Shapes in Inclusive Jet Production in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector*, *Phys. Rev.* **D 83** (2011) 052003 [arXiv:1101.0070] [INSPIRE].

[61] CMS collaboration, *Shape, Transverse Size and Charged Hadron Multiplicity of Jets in pp Collisions at 7 TeV*, *JHEP* **06** (2012) 160 [arXiv:1204.3170] [INSPIRE].

[62] D0 collaboration, *Measurement of color flow in $\mathbf{t\bar{t}}$ events from $\mathbf{p\bar{p}}$ collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV*, *Phys. Rev.* **D 83** (2011) 092002 [arXiv:1101.0648] [INSPIRE].

[63] CDF collaboration, *Study of Substructure of High Transverse Momentum Jets Produced in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev.* **D 85** (2012) 091101 [arXiv:1106.5952] [INSPIRE].

[64] CMS collaboration, *Identification techniques for highly boosted W bosons that decay into hadrons*, *JHEP* **12** (2014) 017 [arXiv:1410.4227] [INSPIRE].

[65] ATLAS collaboration, *Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, *JHEP* **09** (2013) 076 [arXiv:1306.4945] [INSPIRE].

[66] ATLAS collaboration, *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C 76** (2016) 154 [arXiv:1510.05821] [INSPIRE].

[67] J. Gallicchio and M.D. Schwartz, *Quark and Gluon Tagging at the LHC*, *Phys. Rev. Lett.* **107** (2011) 172001 [arXiv:1106.3076] [INSPIRE].

[68] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015 [arXiv:1011.2268] [INSPIRE].

[69] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, *JHEP* **02** (2012) 093 [arXiv:1108.2701] [INSPIRE].

[70] CMS collaboration, *Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS*, 2011 *JINST* **6** P11002 [arXiv:1107.4277] [INSPIRE].

[71] ATLAS collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J.* **C 73** (2013) 2304 [arXiv:1112.6426] [INSPIRE].

[72] ATLAS collaboration, *Measurement of the nuclear modification factor for inclusive jets in Pb+Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV with the ATLAS detector*, *Phys. Lett.* **B 790** (2019) 108 [arXiv:1805.05635] [INSPIRE].

[73] C.W. Fabjan and F. Gianotti, *Calorimetry for particle physics*, *Rev. Mod. Phys.* **75** (2003) 1243 [INSPIRE].

[74] ATLAS collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, *Eur. Phys. J.* **C 73** (2013) 2306 [arXiv:1210.6210] [INSPIRE].

JHEP08(2019)175

[75] CMS collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and MET*, CMS-PAS-PFT-09-001 (2009).

[76] CMS collaboration, *Pileup Removal Algorithms*, CMS-PAS-JME-14-001 (2014).

[77] ATLAS collaboration, *Tagging and suppression of pileup jets*, ATL-PHYS-PUB-2014-001 (2014).

[78] D0 collaboration, *Jet energy scale determination in the D0 experiment*, *Nucl. Instrum. Meth.* **A 763** (2014) 442 [arXiv:1312.6873] [INSPIRE].

[79] ATLAS collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J.* **C 77** (2017) 490 [arXiv:1603.02934] [INSPIRE].

[80] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175] [INSPIRE].

[81] T. Sjöstrand, S. Mrenna and P.Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [arXiv:0710.3820] [INSPIRE].

[82] CTEQ collaboration, *Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions*, *Eur. Phys. J.* **C 12** (2000) 375 [hep-ph/9903282] [INSPIRE].

[83] Shared software for the *Workshop on Mitigation of Pileup Effects at the LHC*, CERN, Geneva Switzerland (2014), https://github.com/PileupWorkshop/2014PileupWorkshop.

[84] *Workshop on Mitigation of Pileup Effects at the LHC*, CERN, Geneva Switzerland (2014), https://indico.cern.ch/event/306155/.

[85] Shared event files for the *Workshop on Mitigation of Pileup Effects at the LHC*, CERN, Geneva Switzerland (2014), http://puws2014.web.cern.ch/puws2014/events/.

JHEP08(2019)175