Karlsruher Institut für Technologie

ETP-KA/2022-08

Dynamic Integration of Heterogeneous Computing Resources and Jet Energy Calibration for the CMS Experiment

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften
(Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

M.Sc. Ralf Florian von Cube

aus Düsseldorf

Tag der mündlichen Prüfung: 4. Februar 2022
Referent: Prof. Dr. Günter Quast
Korreferent: Prof. Dr. Alexander Schmidt

Contents

1. Introduction **1**

2. Computing in High Energy Physics **5**
2.1. The Worldwide LHC Computing Grid 5
2.2. Computing in German High Energy Physics 7

3. Introduction to Quantum Chromodynamics **9**
3.1. The Standard Model of Particle Physics 9
3.2. Quantum Chromodynamics and the Strong Force 11
3.3. Z + Jet Events 12

4. The Large Hadron Collider and the CMS Experiment **15**
4.1. The Large Hadron Collider 15
4.2. The Compact Muon Solenoid 17
4.2.1. Silicon Tracker 18
4.2.2. Electromagnetic Calorimeter 19
4.2.3. Hadronic Calorimeter 19
4.2.4. Solenoid and Iron Return Yoke 19
4.2.5. Muon Chambers 20
4.3. Trigger and Data Acquisition 20
4.4. Event Reconstruction and Particle Flow 21
4.4.1. Particle Flow 22
4.4.2. Pileup 23
4.5. Data Tiers 23
4.5.1. Recorded Data 23
4.5.2. Monte Carlo Simulations 24

5. Jets in CMS and the Jet Energy Calibration **25**
5.1. Jets 25
5.1.1. Jet Clustering Algorithms 26
5.1.2. Pileup Mitigation 27
5.2. Missing Transverse Momentum 28
5.2.1. MET Corrections 28
5.2.2. MET Filters 29
5.3. Jet Energy Calibration 30
5.3.1. Pileup Offset Corrections 31
5.3.2. Simulated Response Corrections 31

5.3.3. Residual Corrections . . . 32
5.4. Absolute Residual Corrections with Z + Jet Events . . . 33
5.4.1. Event Samples and Selection . . . 34
5.4.2. Input Distributions . . . 36
5.5. Derived MPF and p_T-Balance Responses . . . 39
5.6. Global Fitting Procedure and Validation . . . 41

6. Resource Scheduling . . . 47
6.1. Batch Systems . . . 48
6.2. Resource Management with COBALD/TARDIS . . . 50
6.2.1. COBALD . . . 51
6.2.2. TARDIS . . . 52
6.2.3. COVID-19 . . . 53
6.2.4. HTCONDOR Flocking . . . 53
6.3. Site Requirements . . . 53
6.4. Software Deployment and Virtualization . . . 54
6.4.1. CERN Virtual Machine File System . . . 55
6.4.2. Virtual Machines and Containers . . . 55
6.4.3. User Namespaces . . . 57
6.5. Grid Site Operation and Remote Compute Element . . . 57
6.6. Access Rights and Site Policies . . . 58
6.7. Backfilling . . . 60
6.8. Opportunistic Extension of the Aachen Tier 2 WLCG Site . . . 61

7. Conclusion and Outlook . . . 67

A. Appendix . . . 71

List of Figures . . . 75

References . . . 77

1. Introduction

Understanding what matter is composed of and how it interacts has been an everlasting question. The search for the answer drove scientists and philosophers to perform experiments and develop theories to explain their findings. Research in the last centuries has revealed a lot about the constituents of matter and the fundamental forces which govern them.

With the discovery of the electron at the end of the 19th century [1] the search for the smallest constituents of matter gained in momentum. Over the following decades, scientists succeeded in understanding the nucleus of an atom to be composed of protons and neutrons. With the advent of the first particle accelerators, it was possible to gain further insights and validate the theories that had been developed in an attempt to describe experimental observations.

The remaining part of the last century continued to be a period of great success for particle physics. Theories, proven to withstand the scrutiny of experiments, were used to bring an order into the plethora of particles discovered on the way. To continue the journey of understanding matter, research facilities had to be improved constantly. In 1953, the European Organization for Nuclear Research (CERN) was founded for "research of a pure scientific and fundamental character relating to high-energy particles", as is stated in its convention [2]. With a series of particle accelerators, starting with the Synchrocyclotron in 1957 accelerating protons up to an energy of 600 MeV [3], CERN is still driving developments in particle acceleration. This has culminated in the most powerful particle accelerator ever built, the Large Hadron Collider (LHC). However, to keep advancing the field of particle physics, concepts for future machines are continuously being developed.

Today, at the LHC, bunches of protons are accelerated up to a center-of-mass energy of about 13 TeV in two opposing particle beams. Every 25 ns the beams are crossed at multiple points distributed along the circular accelerator, causing collisions of the partons constituting the protons. With large detectors placed at these interaction points, the collisions are recorded to analyze their products. One of these detectors is the CMS detector. It was designed with the goal to test the most successful theory of particle physics, the Standard Model, as well as probe for extensions of it. A description of the LHC and the CMS detector is given in chapter 4.

Collisions at hadron colliders are expected to give rise to numerous partons. Due to their hadronic nature they produce a large number of collimated streams of particles which are called *jets* and interact with the detector material. To get a thorough understanding of the complete underlying physics processes, it is essential to precisely measure the properties of such jets. However, with a lot of different detector components taking part in these complex measurements a discrepancy between the expected and the measured properties of the jets is inevitable. To mitigate these detector effects as well as shortcomings in the predictions, corrections factors can be derived to improve the agreement.

This thesis describes the algorithm used to correct for discrepancies between the measured and modeled jet energy. The correction measurement is a workflow consisting of multiple steps, which are carried out by several groups within the CMS collaboration. The first steps alleviate distortions due to the high frequency at which collisions take place in the detector and multiple protons scattering during one bunch crossing, as well as apparent deficiencies in the simulations. In an additional step, differences between measured and simulated events that remain after applying these first corrections are taken into account. The final correction factors mitigating these differences between measured and simulated responses are derived in this thesis. The so-called *absolute residual corrections* are obtained using recorded data. This is done with events in which a Z boson is produced in association with a jet (Z + jet), stored during the data taking period in year 2018 with the CMS detector.

Two methods are discussed for the derivation of the absolute residual corrections. Both exploit kinematic properties of Z + jet events, these being that the transverse momenta of the Z boson and the jet are in opposite directions. This event feature can be used to calibrate the jet by balancing it against the Z boson. For the analysis, events are selected in which the Z boson decays into a pair of electrons or muons, which can be measured with greater accuracy than the jet. The analysis using these methods and its results is described in chapter 5, after a basic introduction into quantum chromodynamics, the fundamental theory describing quarks and gluons, was given in chapter 3.

HEP experiments such as the CMS detector generate vast amounts of data for physicists to analyze. For such an analysis, the measured low level data is reconstructed to higher level objects and then compared with simulated samples. To generate these simulated samples, the whole process, including the collision of the partons, called hard interaction, the intermediate processes up to the way the products of the collision interact with the detector, and the electronic signals produced by the single detector components are simulated per event. These simulations are based on properties and interactions derived from theoretical models. Computational resources are needed, not only for these event simulations and the reconstruction of the recorded and simulated events, but also for the analysis of the millions of events. Additionally, huge data storage systems to retain all the data, recorded with the detectors, and the simulated samples are required.

The HEP community operates a distributed computing infrastructure to provide these resources: the Worldwide LHC Computing Grid (WLCG). It consists of a sizeable amount of computing centers around the world, which store the data samples and provide computational resources for reconstruction, simulation and analysis. An overview of this computing infrastructure with a special focus on the situation in Germany is given in chapter 2. As discussed in this chapter, this infrastructure is expected to possibly fall short in the medium-term. It is predicted that a lack of computational resources of a factor of about two will arise if no mitigating development is done.

In addition to optimizing simulation and analysis algorithms and advancing or deploying specialized hardware, another approach to address this resource shortage is the allocation of additional non-HEP resources. This approach, however, is not as trivial as it might seem at first glance. The simulations and analyses require specialized software, specific storage systems using community-specific protocols to store samples and analysis results, and possibly access to certain file systems. Resources provided through the WLCG meet these distinct requirements. However, resources from different providers such as university computing clusters or commercial clouds likely do not. Hence, to utilize such resources, additional effort is required to prepare them for transparent usage for HEP workflows.

To provide the required software stacks for analysis and storage access, containerization and virtualization technologies can be used. These have the additional advantage of being able to use already established infrastructure, such as image hubs that are operated commercially or for academia and science, for resource provisioning at different providers.

Another challenge is to use such resources efficiently. Only when HEP resources cannot satisfy the demand should additional external resources be allocated. These allocated resources should be made available in such a way that their use is easy for the user. The implication is twofold: first that resources should be made available dynamically when needed, but also released when this is not the case anymore, and second, that all resources should be provided through one single point of entry.

The integration of such resources is discussed in chapter 6, alongside the resource scheduler COBALD/TARDIS that is being developed at KIT. Different exemplary contexts of usage, especially the successful dynamic integration of an HPC cluster into a WLCG site are demonstrated and issues encountered during this setup process as well as their solutions are discussed.

2. Computing in High Energy Physics

The ever increasing amount of data, recorded by high energy physics experiments needs to be transferred, processed, stored, and analyzed. For this, large infrastructures are deployed and complex algorithms for managing those infrastructures are developed. The infrastructures have to provide the necessary network, compute, and data storage resources to reconstruct raw data, compute data samples for end-user analysis and perform Monte Carlo (MC) simulations of physics processes. Additionally, compute power for end-user analysis is needed to derive physics results. In order to store all raw and intermediate data structures also huge storages are needed. A large fraction of those resources are distributed around the globe as part of the Worldwide LHC Computing Grid (WLCG). The WLCG was planned in parallel with the design of the Large Hadron Collider and its experiments to ensure sufficient resources. Further details on the design of the WLCG are described in section 2.1, and references [4] and [5], which the discussion in section 2.1 is based on.

Many users have additional resources at hand, provided e.g. by their local institutes, or their universities. These resources have the great advantage that they are often available quickly, and close contact with the local operations team of the site helps to easily resolve any issues that may arise. The situation of those resources within the German HEP community is discussed in section 2.2.

2.1. The Worldwide LHC Computing Grid

As mentioned above, the WLCG provides a large amount of computing resources for HEP experiments to perform different tasks. The tiered structure of the WLCG was designed with a very specific top-down mode of operating and distinct functions for the different tiers in mind:

Tier 0 This top-level computing center is located close to the experiments, responsible for first data processing and storage of the raw detector data and the derived data samples. In the original design of the WLCG only one computing center exists at this level.

Tier 1 15 tier 1 computing centers are available for the reprocessing of the data, to produce analysis samples. Furthermore each of the centers archives (tape storage) parts of the raw data, also stored on the tier 0 and provides fast storage (disks) for analysis samples. Tier 1 centers also provide computing resources for analysis

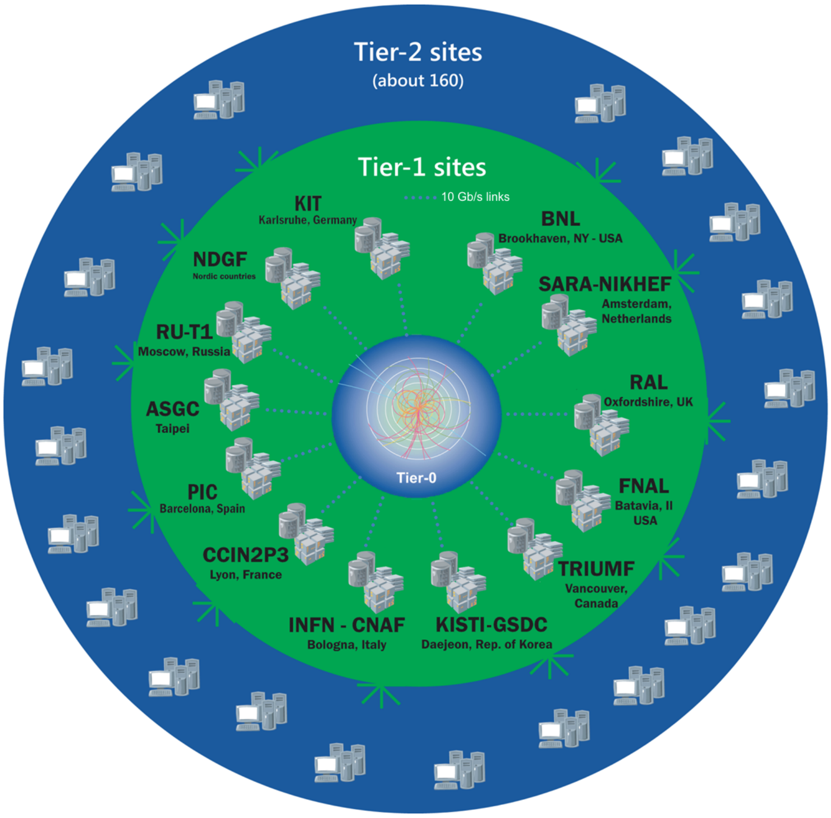

Figure 2.1.: The structure of the WLCG. The tiered design is shown, which implicates the top-down mode of operation of the tasks assigned to the different layers. Figure taken from reference [6]

Tier 2 Additional approximately 170 tier 2 computing centers provide decentralized computing resources for end-user analysis, as well as data storage capacity for those users and the analysis samples required. Those centers are often located at the universities participating in the respective experiments.

This structure is depicted in fig. 2.1. The tier 0 and tier 1 sites, are connected via the dedicated LHC Optical Private Network (LHCOPN). This network connects the majority of the tier 1 sites to the tier 0 with more than 100 GB*p*s. The tier 2 sites in the LHC Open Network Environment (LHCONE) are connected with at least 10 GiB*p*s.

For the upcoming upgrade of the LHC and the experiments, a large increase in need for computing resources is expected. In fig. 2.2 an estimate for CPU time and disk storage requirements by the CMS collaboration up to year 2034 is shown. The CPU requirements are given in kHEP-SPEC06 (kilo HEP Standard Performance Evaluation Corporation 06) [7], the score of a HEP benchmark for CPU performance with the same name. The disk storage requirements are given in Petabyte (PB). The numbers presented are for data processing and analysis, however, only estimate WLCG resources. The blue lines depict two different scenarios, assuming performance improvements in analysis, MC and reconstruction methods and algorithms, as well as improved data formats (solid) and without such developments (dashed). The luminosities per year and trigger rates are adjusted to what is expected in the respective runs. The gray area, enclosed in the black lines, gives a range of the expected

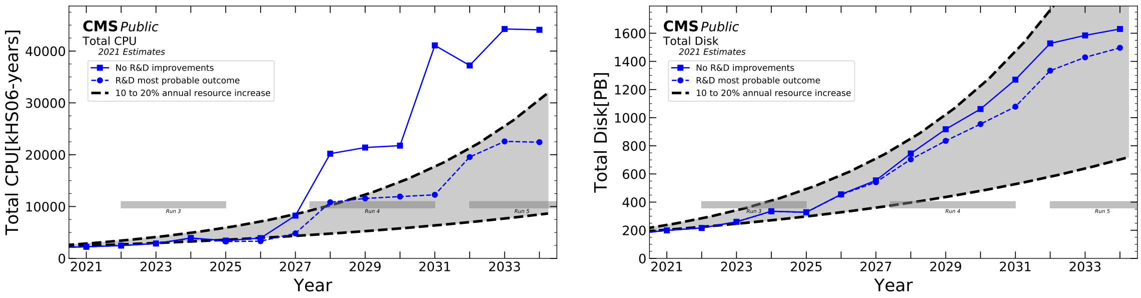

Figure 2.2.: The CPU performance in kHEP-SPEC06 (left) and disk storage in petabyte (right) the CMS collaboration expects to require per year for up to 2034 as of November 2021. Figures taken from reference [8]. In Run 4 a luminosity of 270 fb^{-1} per year with 5 kHz of data saved permanently is expected, and respectively 350 fb^{-1} per year with 7.5 kHz of data saved in Run 5. The solid blue line shows the expected need of WLCG (tier 0, tier 1, and tier 2) resources if neither developments in analysis MC and reconstruction performance nor improvements in data formats are considered. The dashed blue line includes such developments in the projection at their most probable outcome. The black dashed lines with the enclosed gray areas show the expected available resources with only an 10 % to 20 % increase in resources per year. In the best case, no lack of resources is to be expected. In the worst case scenario, a factor of 3 to 4 in terms of computing power and a factor of 2 in terms of disk storage is expected.

resources with an anticipated resources increase of 10 % to 20 % per year. With this model, in the best case enough resources are available to meet the demand. In the worst case, a shortage of a factor of 3 to 4 in terms of computing time and a factor of 2 in terms of disk storage is expected.

2.2. Computing in German High Energy Physics

In German high energy physics the landscape in which computing takes place is very diverse. The physics departments of the Karlsruhe Institute of Technology (KIT), the RWTH Aachen University (RWTH), and the Deutsches Elektronen-Synchrotron (DESY) in collaboration with the Universität Hamburg (UHH) operate official WLCG resources. KIT operates the German tier 1 WLCG site “GridKa”, of which it pledges approximately 5250 cores to the CMS collaboration as of today. RWTH and DESY each operate a WLCG tier 2 site, pledging 2900 and 4600 cores to the CMS collaboration, respectively. Those sites provide for a large amount of compute resources for end-user analysis.

Besides the WLCG resources, the German HEP institutes operate computing resources for their local research groups to provide enough resources for end-user analysis. For all *German* users associated to an LHC experiment, DESY provides the National Analysis Facility (NAF). This site is tightly connected to the WLCG tier 2 computing center, DESY operates. NAF provides users with computing power for analysis with a high-bandwidth connection to the dCache-storage [9] at DESY .

On top of that, each of the institutes also operates its own infrastructures to provide resources for the local groups. For example at the Institute of Experimental Particle Physics (ETP) at

KIT, several local clusters either operated by the institute, the university, or even the federal state are available for, e.g., end-users to perform analysis or generate MC simulations.

This plethora of resources provides the German HEP community with a lot of computing power. However, due to changes in funding, the situation will change in the future and the resource landscape will become more heterogeneous. Funding agencies in Germany will provide computing resources through larger, shared computing clusters for multiple research communities rather than funding individual centers per research community. This will reduce the capacities of German WLCG sites, which has to be counterbalanced by the use of non-WLCG sites. The use of such sites is thus indispensable to keep up the contributions from German sites to the WLCG. A way to consolidate those resources and provide end-users with a transparent way of access, is discussed in chapter 6.

The variety of different kinds of workflows of the HEP community, including, e.g., theory calculations, Monte Carlo (MC) simulations of events, and data analysis, with very different needs, allows HEP computing to take place on a distributed computing infrastructure, comprising heterogeneous resources. Choosing the most fitting site for a given type of workflow, can increase the overall throughput of workflows through the system.

3. Introduction to Quantum Chromodynamics

Scientists always wanted to understand the fundamental processes of phenomena observed in nature as well as experiments and strive to describe them in a unified theory. These theories of our understanding of physics as of today have developed over the last centuries and are yet to be completed. Particle physics, studying the phenomena of the elementary corpuscles forming the matter we interact with every day on a scale only accessible at the highest energies, has come a long way — especially in the last decades.

Particle accelerators like the Large Hadron Collider, accelerating protons nearly up to the speed of light, prior to bringing them to collision within experiments like the CMS detector, are testing the understanding of the theoretical models of particle physics we have today. Based on measurements of such machines, the theories can be tested, refined, and free parameters of the models can be determined.

3.1. The Standard Model of Particle Physics

The objective of particle physics is to develop a unified theory of the interaction of the elementary particles, describing their kinematic properties and the forces they exert onto each other. This theory is called the Standard Model (SM) of particle physics. It comprises the description of the fundamental forces: the electromagnetic, the weak, and the strong force. The fourth fundamental interaction, although mathematically described for centuries, the gravitation, could not yet be included in the unified description of interactions, which the SM represents. However, very weak on the scales of hadron collider physics, it can be neglected in the following discussions. This section is very loosely based on reference [1, chapter 1].

The forces contained in the picture of the SM are described by their underlying fundamental theories: the strong force emerges from quantum chromodynamics (QCD), the electromagnetic force and the weak force are combined within the electroweak theory (EW). The SM accomplishes a unification of those theories, however, leaves a few parameters to be determined by experiments. These parameters include, but are not limited to, certain masses of the elementary particles and the coupling strength parameters of the forces.

In addition to the forces, the SM describes different types of particles: the fermions contain quarks and leptons which make up matter, and the class of bosons comprising the gauge bosons which mediate the forces and the Higgs boson, closely associated to the mechanism responsible for the masses of the particles. Quarks exist with two different electric charges: up (u), charm (c), and top (t) carry electric charge $+\frac{2}{3}$ in units of elementary charge. Those are referred to as *up-type*. They are complemented by the *down-type* quarks down (d), strange (s), and bottom (b) which carry electric charge $-\frac{1}{3}$. Additionally the quarks carry a weak charge which is usually described by the third component of the weak isospin, T_3. The *up-type* quarks carry $T_3 = +\frac{1}{2}$, while *down-type* quarks have $T_3 = -\frac{1}{2}$. To each quark its corresponding antiparticle, the anti-quark, exists, which carries opposite charges, but all other quantum numbers remain the same. Furthermore, each quark carries a color *red*, *green*, *blue*, or the corresponding anti-color *anti-red*, *anti-green*, or *anti-blue*. The color is the charge of the strong force, which will be discussed below. Quarks therefore participate in electromagnetic, weak, and strong interactions.

Leptons can be organized in two groups: *electron-like* leptons, i.e. the electron (e), muon (μ), and tau (τ), in order of increasing mass, carry electric charge -1 and weak isospin of $T_3 = -\frac{1}{2}$, while their corresponding neutrinos (ν_e, ν_μ, and ν_τ) are electrically uncharged

Table 3.1.: The particles of the Standard Model of particle physics. The column "Forces" indicates which interactions the particles participate in with "E", "W", and "S" encoding electromagnetic, weak and strong interaction, respectively. For the gauge bosons, the interaction they mediate is given in the brackets after their name. The charge of the particles is given in units of the elementary charge. Masses are taken from reference [10].

			Name	Charge / [e]	Mass	Forces
Fermions	Quarks	u	up	$2/3$	2.16 MeV	E, W, S
		d	down	$-1/3$	4.67 MeV	E, W, S
		c	charm	$2/3$	1.27 GeV	E, W, S
		s	strange	$-1/3$	93 MeV	E, W, S
		t	top	$2/3$	172.8 GeV	E, W, S
		b	bottom	$-1/3$	4.18 GeV	E, W, S
	Leptons	e	electron	-1	511 keV	E, W
		ν_e	electron neutrino	0		W
		μ	muon	-1	105.7 MeV	E, W
		ν_μ	muon neutrino	0		W
		τ	tau	-1	1.777 GeV	E, W
		ν_τ	tau neutrino	0		W
Bosons	gauge	g	gluon (S)	0		S
		γ	photon (E)	0		
		Z	Z boson (W)	0	91.19 GeV	W
		$W^\pm$	W boson (W)	± 1	80.38 GeV	E, W
		H	Higgs boson	0	125.1 GeV	W

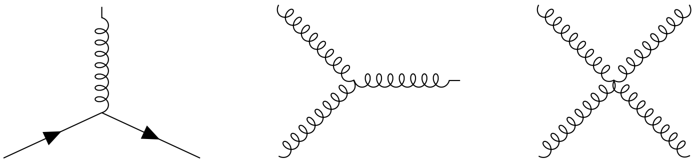

Figure 3.1.: The fundamental vertices of quantum chromodynamics (QCD): The interaction of a quark (—▶—) with a gluon (∿∿∿) on the left. The two right diagrams portray the triple and quartic self-coupling vertices of the gluon due to its color-charge.

and carry a weak isospin of $T_3 = +\frac{1}{2}$. Charged leptons participate in electromagnetic and weak interactions, neutrinos only in weak interactions.

The quarks and leptons interact by the exchange of force mediating gauge bosons. The gluons (g) are the mediators of the strong force which also carry a color charge leading to self-interactions which is discussed below and shown in fig. 3.1. The electromagnetic force is mediated by the photon (γ), while the electrically neutral Z boson and the positively and negatively charged W bosons transmit the weak force. The bosons are a direct consequence of the requirement of invariance under local gauge transformations of the respective fields of the underlying quantum field theories.

The last particle, the Higgs boson, is a result of the Higgs mechanism, the theoretical description of how the massive particles of the SM acquire their mass.[1] An overview of the particles and some of their properties are summarized in table 3.1.

3.2. Quantum Chromodynamics and the Strong Force

Quantum chromodynamics (QCD) is the theory of the strong interaction which quarks and gluons are governed by. When bringing protons to collision at the LHC, the interacting particles are the quarks and gluons, which the proton consists of. Each of these quarks and gluons carry a fraction of the momentum of the proton. This fraction influences the probability of them taking part in the interaction. The probability is described by the parton distribution function (PDF), as discussed in more detail in reference [12].

The following discussion is based on references [1, chapters 2 and 8] and [13].

The fundamental interaction vertices of the strong interaction are given in fig. 3.1. These are combined to derive more complex processes. Each vertex in the Feynman diagrams introduces a factor proportional to the gauge coupling constant $g_s = \sqrt{4\pi\alpha_s}$ of the strong interaction. The “constant” depends on the momentum scale μ^2. For increasing μ^2, the coupling constant decreases. This means, at high energies and small distances, quarks and gluons can be approximated as free particles in perturbation theory. This is known as *asymptotic freedom*, which was introduced in reference [14]. At lower energies and

[1] An exception are neutrinos, which are massless in the current state of the SM, but are observed to oscillate between their mass eigenstates, implying a non-zero mass. [11]

larger distances, on the other hand, the coupling $\sqrt{\alpha_s}$ increases, as the energy in the field of the strong interaction increases. This results in particles being created when spatially separating color-charges, as it reduces the potential energy in the strong field. This is known as *confinement* and as a consequence, only bound states of color-charged particles can exist freely in nature.

As described above, the charge, strongly interacting particles carry, is called color. Each quark carries either *red*, *blue* or *green* as color, accordingly anti-quarks carry a unit of anti-color *anti-red*, *anti-blue* or *anti-green*. Gluons carry color and anti-color. The theory, based on the $SU(3)$ group with $T_a = \lambda_a/2$, where λ_a are the Gell-Mann matrices, as its generators, predicts a color singlet and a color octet, as introduced in reference [15]. Each of the octet states corresponds to a gluon, which, due to the net color charge they carry, couple to each other, causing the triple and quartic self-coupling, shown in fig. 3.1.

Because of the asymptotic freedom in deep inelastic scattering experiments at high energies as e.g. the experiments at the Large Hadron Collider, the hard scattering process can be approximated as an interaction of free quarks and gluons, also called *hard interaction*. The colliding particles are the *initial-state* particles, the products of the hard interaction the *final-state* particles. Due to their high energies, strongly interacting final state particles might radiate off gluons. This process can repeat, causing numerous color-charged particles to be radiated off, generating a *parton shower* and losing energy during this process. Reaching a certain energy scale, the showering stops and instead a cascade of color-neutral particles, so-called hadrons, is produced. This process is called *hadronization*. The particles generated during these processes are eventually measured by the detector. The process from the initiating final state particle, undergoing parton-showering and hadronization gives rise to collimated streams of particles, which are called *jets*. Jets and the calibration of the measurement of their energy are described in detail in chapter 5.

3.3. Z + Jet Events

For the calibration of the jet energy scale, as will be discussed in chapter 5, events in which a Z boson is produced in association with a jet are used. The Z boson, discovered in 1983 [16], is a prime candidate for calibration purposes at the LHC, as its properties have been studied in great detail in the past. Its mass

$$M_Z = 91.1876 \pm 0.0021\text{GeV}$$

and full decay width

$$\Gamma = 2.4952 \pm 0.0023\text{GeV}$$

have been measured with great precision [10] and are often used for checking data consistency in analyses.

For the calibration, events are used in which a Z boson is produced in association with a jet, as the Z boson itself can be reconstructed rather precisely from a pair of electrons or muons. In QCD theory calculations, complex problems are often not analytically solvable

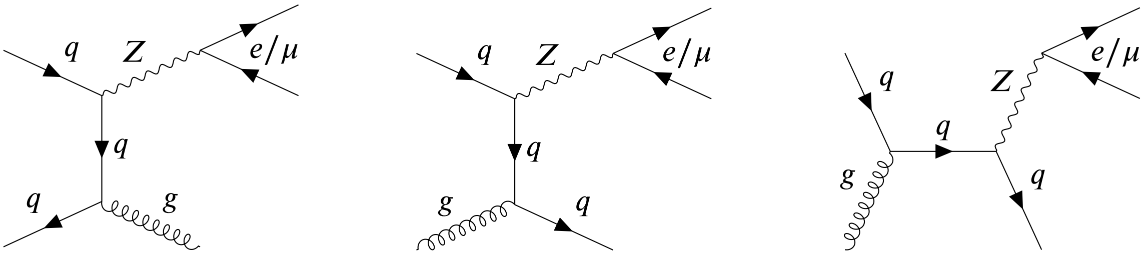

Figure 3.2.: The Feynman diagrams of the leading order production processes of a Z boson (∿∿∿) in association with a quark or gluon and the Z boson decaying into a pair of electrons or muons. The left diagram depicts the quark-antiquark production channel, the center diagram the quark-gluon channel via the exchange of a quark and the right diagram the quark-gluon channel via an intermediate quark. The strongly interacting final state particle produces a jet of particles which can be detected in the detector material.

to complete precision, hence processes are Taylor expanded in orders of $\sqrt{\alpha_s}$. This is called perturbative QCD and is discussed in greater detail in reference [13].

Expanding the process up to only leading order, the Z + jet production happens either in a quark-antiquark channel, a quark-gluon channel via the exchange of a quark or a quark-gluon channel via an intermediate quark. These production processes, with the Z boson afterwards decaying into a pair of electrons or muons, are shown in fig. 3.2.

When producing such events at the LHC, the initial state particles collide head on. Their momentum transverse to the beam line is negligible. Thus, due to momentum conservation, the sum of the transverse momenta of the final state particles has to vanish as well. This allows to balance the measured jet against the Z boson, using the latter as reference object. This event property will be indispensable in the derivation of the jet energy correction, discussed in chapter 5.

4. The Large Hadron Collider and the CMS Experiment

To learn about the nature of particles, large facilities are built to investigate their properties. Particles are accelerated up to nearly the speed of light before being brought to collision. During this process, new particles are generated that traverse elaborate detectors built around the collision point.

By moving through the detector and interacting with its material, the particles leave tracks and energy deposits in different parts of the detector. Those traces are recorded electronically and made accessible for physicists. Using sophisticated reconstruction algorithms and analysis techniques, analyses of the recorded signals are performed and high-level variables for comparison with theory are derived and measured.

In the following sections, the processes of how protons are accelerated and brought to collision, how the generated particles leave traces in the detector and how those traces, manifesting as electronic signals, are used to derive particle trajectories, are discussed. The Large Hadron Collider with its accelerator complex is described in section 4.1. In the subsequent section 4.2, the different components of the Compact Muon Solenoid detector are described and how the measurements of the different subdetectors are combined is outlined.

4.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is a particle accelerator located underground at the border between Switzerland and France, next to Lake Geneva. Approximately 100 m underground in the tunnel of its predecessor, the Large Electron-Positron Collider, protons and ions are brought up to nearly the speed of light through a series of accelerators. The LHC is the final and largest machine in this acceleration chain and as of today the machine reaching the highest center-of-mass energy. Protons traverse the ring, running in two beams in clockwise and counterclockwise directions. These two beams are brought to collision at four interaction points (IPs). At those IPs the detectors are located, recording the particles produced in the collision.

The discussion in this section is based on reference [17], which can be referred to for further details.

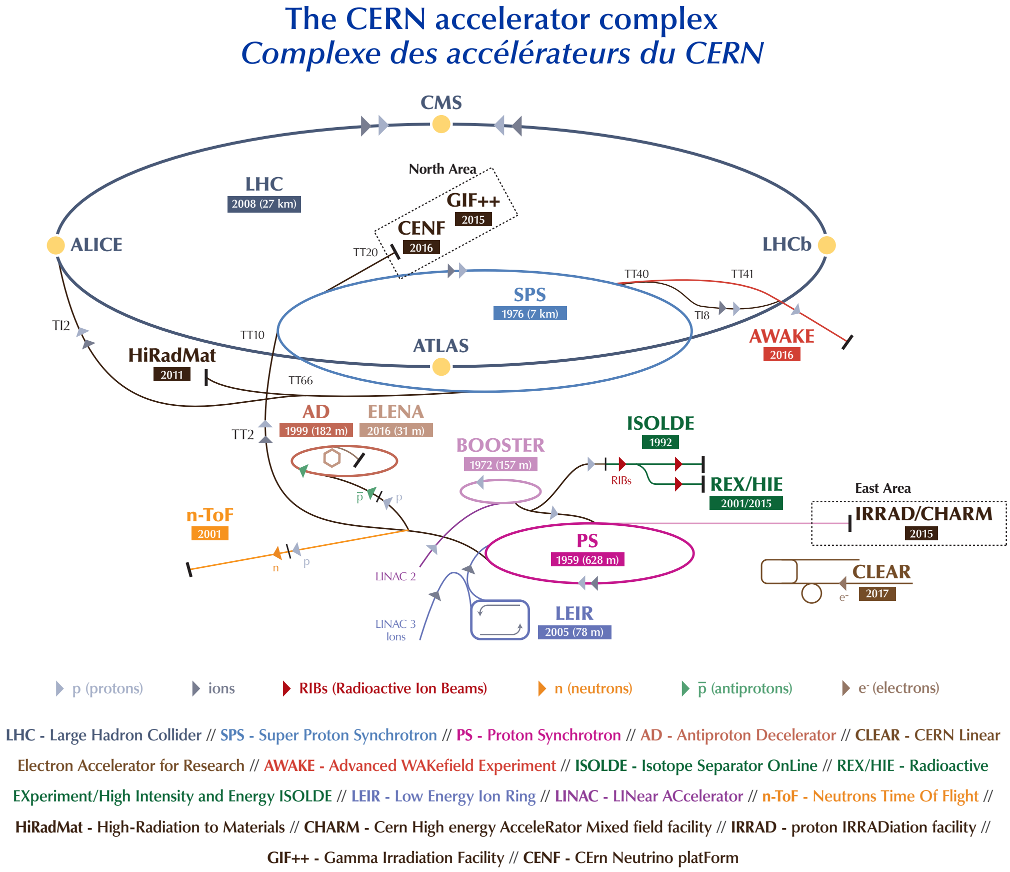
The CERN accelerator complex
Complexe des accélérateurs du CERN
CMS
North Area
GIF++
2015
CENF
2016
LHC
2008 (27 km)
ALICE
LHCb
TT20
TT40
TT41
SPS
1976 (7 km)
TI8
TI2
TT10
ATLAS
AWAKE
2016
HiRadMat
2011
TT66
AD
1999 (182 m)
ELENA
2016 (31 m)
TT2
BOOSTER
1972 (157 m)
ISOLDE
1992
RIBs
REX/HIE
2001/2015
East Area
IRRAD/CHARM
2015
n-ToF
2001
p̄
p
n
p
PS
1959 (628 m)
CLEAR
2017
e⁻
LINAC 2
LINAC 3
Ions
LEIR
2005 (78 m)
p (protons)
ions
RIBs (Radioactive Ion Beams)
n (neutrons)
p̄ (antiprotons)
e⁻ (electrons)
LHC - Large Hadron Collider // SPS - Super Proton Synchrotron // PS - Proton Synchrotron // AD - Antiproton Decelerator // CLEAR - CERN Linear
Electron Accelerator for Research // AWAKE - Advanced WAKefield Experiment // ISOLDE - Isotope Separator OnLine // REX/HIE - Radioactive
EXperiment/High Intensity and Energy ISOLDE // LEIR - Low Energy Ion Ring // LINAC - LINear ACcelerator // n-ToF - Neutrons Time Of Flight //
HiRadMat - High-Radiation to Materials // CHARM - Cern High energy AcceleRator Mixed field facility // IRRAD - proton IRRADiation facility //
GIF++ - Gamma Irradiation Facility // CENF - CErn Neutrino platForm

Figure 4.1.: The Large Hadron Collider accelerator complex. Figure taken from reference [18].

The proton-proton collisions at the LHC take place at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. The acceleration of the particles is done in multiple steps. The protons have their source in a bottle of hydrogen gas installed before they are accelerated for the first time in the linear accelerator LINAC 2 (shown in purple in the lower region of fig. 4.1). With an energy of 450 MeV, the protons reach the Proton Synchrotron Booster (BOOSTER), the first circular accelerator. Here, the acceleration continues until the protons reach an energy of 1.4 GeV. The next stage is the Proton Synchrotron (PS), where an energy of 25 GeV is reached, followed by the last pre-accelerator, the Super Proton Synchrotron (SPS). With an energy of 450 GeV, the protons are then fed into the LHC.

Due to the nature of the synchrotron pre-accelerators, the protons enter the LHC in bunches. With a distance of 25 ns, these "packets" of approximately 1.15×10^{11} protons circulate inside the ring. In this last stage, the protons are accelerated to reach their nominal energy of 6.5 TeV. With the geometry of the bunches at the IPs, this yields an instantaneous luminosity of approximately

$$L = 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}.$$

At eight locations around the ring, the LHC can be accessed from above ground. At each point either experiment or beam monitoring and maintenance infrastructure is installed. The

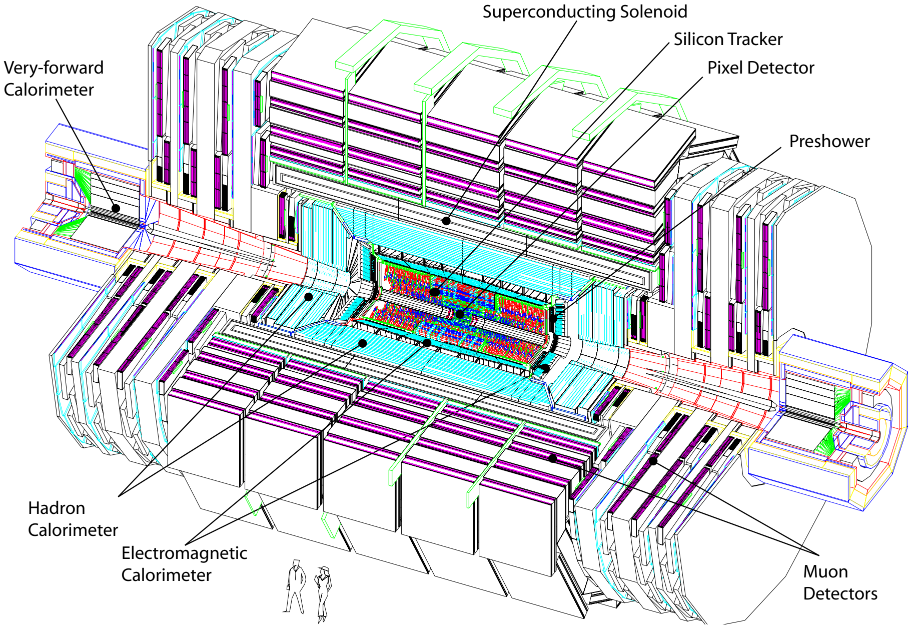

Figure 4.2.: The CMS experiment with its subdetectors. Figure taken from reference [21]

four major experiments ALICE, ATLAS, CMS, and LHCb are marked yellow in fig. 4.1. ATLAS and CMS are multipurpose experiments designed to investigate a wide range of physical processes, while ALICE focuses on studying quark-gluon plasma and LHCb with its fundamentally different, asymmetric setup concentrates on decays containing bottom and charm hadrons. Not shown in fig. 4.1 are LHCf (Large Hadron Collider forward) and TOTEM (Total Elastic and Diffractive Cross Section Measurement). The former is located at 140 m from the ATLAS detector between the two beam pipes and can be used for luminosity measurements, the latter is in the vicinity of the CMS detector and studies the total proton-proton cross-section, as well as elastic and diffractive scattering. [19, 20]

4.2. The Compact Muon Solenoid

One of the multipurpose detectors is the Compact Muon Solenoid (CMS) experiment. It is located at interaction point 5 of the LHC accelerator complex. The CMS detector, as shown in the sketch in fig. 4.2, has a cylindrical layout, centered around the interaction point. In a layered structure, different detector technologies are used to record the trajectories of the particles traversing the detector.

At CMS, a Cartesian coordinate system that aligns with the nominal interaction point is used. It is outlined in fig. 4.3. The x-axis points towards the center of the LHC, the y-axis upwards, and the z-axis in the direction of the counterclockwise beam. Furthermore, a spherical coordinate system is used, also shown in the figure, reflecting the cylindrical symmetry of the CMS detector: The azimuthal angle, φ, is measured from the x-axis in the plane transverse to the beam. The polar angle, ϑ, is measured from the z-axis. To describe

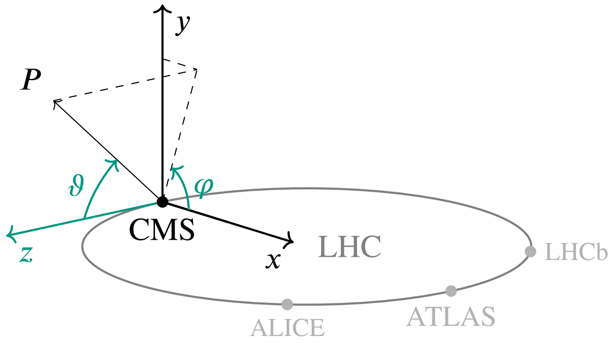

Figure 4.3.: The coordinate system used at CMS. The x-axis points to the center of the LHC, the y-axis points upwards, and the z-axis in counterclockwise direction. The azimuthal angle φ is measured in the plane transverse to the beams, the polar angle ϑ is measured from the z-axis. Figure adapted from reference [22].

events in the detector, most often the pseudorapidity $\eta = -\ln(\tan(\vartheta/2))$ is used instead of the polar angle ϑ. This is done because for highly relativistic particles pseudorapidity differences are invariant under the Lorentz transformation, meaning they do not depend on the longitudinal boost of the reference frame.

In the following, a short introduction to each subdetector is given. The descriptions follow roughly reference [21], which can be consulted for a more thorough discussion.

4.2.1. Silicon Tracker

At the center of the layered detector design, directly around the interaction point, is the inner tracking system. It is designed to precisely measure charged particles, especially their momentum, and the position of the collision in the detector, taking advantage of the Lorentz force bending the trajectories of charged particles in magnetic fields.

The innermost part is the *pixel detector*. It consists of four cylindrical layers (BPix) with a length of 54 cm comprising silicon pixels with a size of 100 µm × 150 µm at radii 2.9 cm, 6.8 cm, 10.9 cm, and 16.0 cm centered around the interaction point. With six disks of pixels (FPix), each covering the radius range from 4.5 cm up to 16.1 cm, mounted at $z = \pm 29.1$ cm, $z = \pm 39.6$ cm, and $z = \pm 51.6$ cm complementing the BPix on either end, the pixel tracker covers a pseudorapidity range of $|\eta| < 2.5$. [23]

The outer part of the silicon tracker is the *silicon strip tracker*. This part of the silicon tracker covers the radial range from 10 cm to 116 cm and, like the pixel detector, the pseudorapidity range $|\eta| < 2.5$. In four different parts, the tracker inner barrel (TIB) and disk (TID), the tracker outer barrel (TOB), and tracker end caps (TEC), different kind of silicon strip sensors are used. The whole tracking system is read out with $\sim 75 \times 10^6$ channels.

4.2.2. Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is the next layer surrounding the tracker. It aims to precisely measure the energy of electrons and photons. To measure the energy of electromagnetic particles, lead tungstate ($PbWO_4$) crystals are used, which absorb the particles and convert their energy into photons, which in turn are collected and read out by photodetectors. The material $PbWO_4$ has a very high density of $8.28\,\mathrm{g\,cm^{-3}}$, resulting in a short radiation length of 0.89 cm. In combination with the small Molière radius of 2.2 cm, this allows for a compact and highly granular design of the ECAL.

The ECAL is structured in various subsystems using distinct crystal shapes and different photodetectors. The ECAL barrel (EB) covers the pseudorapidity range $|\eta| < 1.479$ within the radius range of 129 cm to 152 cm. The length of this geometry corresponds to 25.8 radiation lengths. The ECAL endcaps (EEs) cover the pseudorapidity range $1.479 < |\eta| < 3.0$. They are placed at $z = \pm 315.4\,\mathrm{cm}$ with a crystal length of 220 mm corresponding to 24.7 radiation lengths. The ECAL preshowers (ESs) are located in front of the EEs. The ESs are sampling calorimeters, combining lead absorber plates with silicon strip sensors.

The whole ECAL consists of 75 848 lead tungstate crystals covering the pseudorapidity range up to $|\eta| < 3.0$, each with an individual readout channel.

4.2.3. Hadronic Calorimeter

The hadronic calorimeter (HCAL) is located outside the ECAL. Its purpose is to measure the energy of particles not completely depositing their energy in the ECAL. The HCAL as a sampling calorimeter measures energies by stacking absorber material with active material, in this case brass plates and plastic scintillators respectively. After being collected by wavelength-shifting fibers, the scintillation light yield is read out using silicon photomultipliers.

Like the ECAL, the HCAL is also composed of various subsystems covering different pseudorapidity and radius regions. The HCAL barrel (HB) covers the region $|\eta| < 1.3$ in the radius range from 1.77 m to 2.95 m, right up to the solenoid. On the outer side of the solenoid, the HCAL outer (HO), also called *tail catcher*, covers the same pseudorapidity range, using the solenoid as an absorber material. Located next to the EEs, the HCAL endcaps (HEs) cover the region $1.3 < |\eta| < 3.0$. In the forward region $2.8 < |\eta| < 5.2$, located at 11.2 m from the IP, the HCAL forwards (HFs) use Cherenkov light from the electromagnetic component of showers to measure the energy.

4.2.4. Solenoid and Iron Return Yoke

One of the defining features of the CMS detector is the superconducting solenoid, generating a homogeneous 4 T magnetic field inside. The 220 t solenoid measures 6 m in diameter by 12.5 m in length. It is powered by a 20 kA power supply. In its free bore, it contains the tracking and calorimetry systems discussed above. A 10 000 t iron return yoke, which confines the magnetic field to the detector volume, provides structural support for the vacuum vessel, containing the four layers of the niobium-titanium superconductor. The solenoid and the return yoke together make up approximately 90 % of the detector's mass.

Because of the magnetic field, charged particle trajectories, adhering to the Lorentz force, are bent, which contributes to their identification. Based on the direction of deflection of the trajectory of a charged particle, its charge sign can be determined. The curvature of the particle's track can be used for measuring its momentum, which is an essential part of the energy measurement. This is especially important for muons, that are not fully absorbed by the calorimeters but traverse the detector to its most outer regions.

4.2.5. Muon Chambers

Embedded within the iron yoke is the muon system. Muon detection at CMS is performed using different types of gaseous detectors. 250 drift tube (DT) chambers in four layers cover the pseudorapidity range of $|\eta| < 1.2$. Charged particles passing through the argon-carbon dioxide filled DTs ionize the gas. The free electrons move to the positively charged sensitive wires and produce a signal. By combining signals from multiple tubes, the trajectory of the particle can be reconstructed. In the endcap region, 468 cathode strip chambers (CSCs) cover the pseudorapidity range up to $|\eta| < 2.4$. The use of anode wires with perpendicular cathode strips allow for precise measurements, even with inhomogeneous magnetic fields. Complementing the DTs and the CSCs, resistive plate chambers (RPCs) are installed up to $|\eta| < 1.6$. These consist of two resistive, parallel plates separated by gas. Again, traversing particles ionize the gas and the resulting avalanche of electrons produces a signal in the detecting strips. From the signal of the electrons on the strip, a good estimate of the primary particle momentum can be made. As the response times of the RPCs are well below the 25 ns between two bunch crossings [24], their signals can be used for triggering, cf. section 4.3.

4.3. Trigger and Data Acquisition

With a bunch crossing frequency of 40 MHz and on average 20 simultaneous proton collisions at each crossing, even more for later run periods, at the given luminosity, it is impossible to record all events. In order to be able to reduce the number of events permanently stored while not discarding events of great interest, a two-tiered trigger system is installed. This discussion loosely follows reference [25].

The level 1 trigger (L1) is implemented in custom, programmable hardware. It aims to reduce the data output to a rate of up to 100 kHz with a latency of 4 µs. For the L1, low level detector readouts are used to decide whether to drop an event or forward it for further reconstruction. These decisions are based on *trigger primitives* which are first generated using energy deposits in the calorimeters (ECAL and HCAL), or track segments and hits in the muon system (DTs, CSCs, and RPCs).

Trigger primitives from the calorimeters consist of the transverse energies and quality flags from the calorimeter towers. They are then combined by the regional calorimeter trigger (RCT) to trigger objects. Such object can be an electron or photon candidate or regional transverse energy. The trigger objects are sorted and assessed by the global calorimeter trigger (GCT), which also finds jets, calculates the missing transverse energy, and then outputs the global calorimeter trigger objects.

Trigger primitives originating from the muon subsystems DTs, CSCs, and RPCs are combined by their respective track finders. These tracks are then transmitted to the global muon trigger (GMT), which combines candidates with tracks in multiple subsystems to avoid double counting and assesses the quality of the muon candidates. The GMT outputs global muon trigger objects.

The final decision of whether to keep an event at this stage is implemented in L1 trigger menus and performed by the L1 global trigger system (GT). A menu contains up to 128 “paths”, which are evaluated in parallel. Each path acts on the global trigger objects and incorporates technical and external triggers, which include other signals from the CMS detector. All paths are combined with a logical OR, resulting in the final decision of the GT. If this decision is positive and all systems are available, the trigger control system (TCS) fires an L1 accept (L1A) signal. The decision process happens in less than 1 µs.

On L1A, the detector readout is transferred from the front-end buffers through the data acquisition (DAQ) to the high level trigger (HLT). The DAQ has to withstand an input rate of 100 kHz from the various subdetector components, the maximum output rate of L1. The HLT performs a much more complex analysis compared to the L1, running a simplified reconstruction workflow on the complete detector readout, including information from the various detector subsystems, and reduces the number of events, which will finally be stored offline by a factor of 1000. While L1 is hardware based, the HLT is implemented in software and runs on a compute-farm. Therefore, more complex analyses can be performed on the detector readout and the algorithms can be adapted as needed. To keep trigger efficiencies high, the reconstruction algorithms are close to the ones applied afterwards in offline computing. The output rate of the HLT is approximately 100 Hz with an event size of roughly 1.5 MB. [26]

4.4. Event Reconstruction and Particle Flow

The raw data measured by the different subdetectors is not suitable directly for end-user analysis. For this, the low-level detector entries, consisting e.g. of measurements of energy deposits in the calorimeters or tracker hits, need to be converted into more high-level objects, like particles, jets, or event quantities like the missing transverse energy. At CMS, this reconstruction of particle candidates is done by the particle flow (PF) algorithm. It is described in more detail in reference [27], on which also the following discussion is based. The PF algorithm starts by reconstructing basic elements from the detector entries:

Charged Particle Trajectories and Vertices Trajectories of charged particles and vertices are reconstructed using tracker hits in an iterative procedure. A number of neighboring tracker hits are combined and used as seeding hits for a fitting procedure, optimizing with respect i.a. the resulting compatibility with the primary vertex. The hits associated to the track are removed from the collection and from the remaining hits a new fitting procedure with looser quality criteria is started. Well-performing tracking of charged particles is crucial, as for example an untracked charged hadron would be reconstructed with the calorimeters as neutral and therefore heavily distort the energy fractions used for jet identification.

Calorimeter Clusters Another element for the generation of the PF candidates are the calorimeter clusters, used to measure the energy of particles, identify electrons and photons, and improve the energy measurements of charged hadrons. In the ECAL, a cluster seed consists of a single lead tungstate crystal, while in the HCAL multiple crystals form a cluster seed. The clustering in the multiple subsystems of the ECAL and HCAL is done separately with the exception of the HF, in which the energy deposits are not clustered.

The clustering starts from cluster seeds above a certain threshold energy and with higher energies than the cells in their vicinity, hence a local energy maximum. By an iterative algorithm, cells above the threshold, sharing a corner with the seed are added, and topological clusters are built. If a cell can be associated to several topological clusters, the energy is split between them, depending on the distance. Afterwards, assuming Gaussian energy deposits of the seeds within the topological cluster, the spatial position and energy of the clusters are determined by fitting a Gaussian mixture model with a maximum-likelihood fit.

Muon Tracks A similar approach as for the reconstruction of the tracks is used for the reconstruction of muons. Three different types of muons are considered:

1. *Standalone Muons* are reconstructed starting from hits in the DTs and CSCs as fitting seeds. For the final fit, the hits in RPCs are also taken into account.
2. When a standalone muon can be matched with a compatible track from the tracking system, these hits get combined for the fitting procedure, resulting in a *global muon*.
3. Each track in the tracking system above a certain momentum threshold is extrapolated to the muon system. If a compatible hit is found in the muon system, satisfying specific requirements, the track in the tracking system is considered a *tracker muon*.

4.4.1. Particle Flow

A particle in the detector usually creates multiple PF elements such as tracks and calorimeter clusters. To generate the particle, those elements from the different subdetectors are combined with a *link algorithm*. The algorithm combines several PF elements to *PF blocks* depending on their distance within the (η, φ)- or (x, y)-plane and specific rules for the different elements. The link can either be directly between two elements or chained with another common element between them. To generate the reconstructed particle, the PF blocks are then combined based on the type of particle candidate. All elements are then removed from the collection for the next particle reconstruction.

The link algorithm results in particles, each one of type charged hadron, electron, muon, photon, or neutral hadron. [28] These are saved as PF candidates which subsequently pass through further reconstruction algorithms to determine higher-level objects like the missing transverse momentum or jets, as discussed later in section 5.1.1.

4.4.2. Pileup

The PF algorithm assumes all reconstructed particles, jets, and other objects to originate from the same hard interaction. However, as protons at the LHC are accelerated and brought to collision in bunches, as discussed in section 4.1, multiple scatterings during the same bunch crossing are inevitable. These additional scatterings cause tracks and energy deposits in the detector and contaminate the detector readout, assumed to be from the hard interaction. Such additional detector activity within one bunch crossing is called *in-time* pileup (PU). Accordingly, *out-of-time* PU is activity in the event, which originates from the previous or the next bunch crossing. The former can happen if the detector is read out before the calorimeter is “cooled down”, meaning the signal present from the previous bunch crossing is not yet depleted. The latter happens, if the integration time of the measurement is too long, continuing the measurement when the next bunch crossing already occurs.

Such effects have to be taken into account in the later analysis of the events. Concerning jets, such algorithms are discussed in section 5.1.2. As the clustering and the reconstruction of the jets, however, happens before the analysis, jets have to be corrected for the offset in their energy due to PU, as will be discussed in section 5.3.1.

4.5. Data Tiers

Events are stored as collections of their particles and associated information, following the Event Data Model (EDM) in `root`-files [29], a compressed file-format which provides tree-like access to the contents of the events. With several successive data tiers, starting from low-level detector readout to reconstructed and calibrated physics objects ready for high-level analysis, different degrees of detail are provided to perform physics studies. Owing to the different amounts of information contained, the data tiers vary widely in their size. The following discussion is based on reference [26].

4.5.1. Recorded Data

The detector readout after passing through the L1 and the HLT is called RAW. As mentioned before, in this data tier, each event has a size of approximately 1.5 MB. It is used as input for the reconstruction, which produces the RECO data tier as output. In the latter, roughly 0.25 MB are stored per event, containing very detailed information of the included objects. For physics data analysis, the Analysis Object Data (AOD) data tier is derived. It also contains the reconstructed objects, but the degree of detail of the information is decreased. Because of this, the size can be reduced to 0.05 MB per event. AOD is the first data tier viable for use in physics analysis.

Even more reduced data tiers called MINIAOD and NANOAOD, in order of decreasing storage size per event, are derived from AOD and MINIAOD, respectively. The contained objects store less detailed reconstruction information but are already highly optimized for standard physics analyses. This means, particles are already corrected for all known detector and reconstruction inefficiencies as known at the time of production of the files, or at least the correction factors are provided with the events. Preparing the data this way reduces the time needed to process an event in the final analysis but especially decreases the amount of data that needs to be transferred from storage to the machine running the analysis.

However, as information is heavily condensed, specific analyses might require additional information on specific objects that are not provided in the basic NANOAOD. This is approached by providing analysis group specific versions of the NANOAOD data tier, which are augmented with the necessary information.

4.5.2. Monte Carlo Simulations

Events are simulated starting from the fundamental inelastic scattering of the protons. The final state particles emerging from the hard interaction undergo parton-showering, hadronization, and underlying event simulations using e.g. PYTHIA8 [30] or HERWIG [31, 32]. The result is referred to as *particle-level*. The last stage also including the simulation of the detector's response using e.g. GEANT4 [33], called *detector-level*, completes the data tier GENSIM. Additional pileup collisions following the expected pileup distribution are simulated and overlayed. This is the lowest level data tier concerning MC simulated data. The different levels in GENSIM are crucial in gaining a better understanding of the detector response by comparing them to later stages in reconstruction.

Now, at a stage corresponding to the raw detector readout, reconstruction can take place as with RAW when handling recorded data. The output is appropriately referred to as RECOSIM, the postfix "SIM" indicating that the data is simulated. Subsequent data tiers are derived, adhering to the same data flow of reconstruction algorithms and data reduction as above, while suffixing the names with "SIM".

5. Jets in CMS and the Jet Energy Calibration

As discussed in chapter 3, the interacting particles at the LHC are not the protons themselves but rather their constituents, the quarks and gluons. Owing to their nature of interacting strongly, in combination with the high energies at which collisions take place at the CMS experiment, quarks and gluons hadronize close to the interaction point (IP) and generate numerous “streams of particles” that traverse the detector. These streams are called jets and can be used to study the underlying processes. They often originate from the hard interaction process, including also radiative processes before or after the collision.

In this chapter, the reconstruction of such jets, starting from PF candidates as introduced in section 4.4, is described. As these jets, however, are highly intricate and measurements from different detector components are combined for their reconstruction, great care has to be taken to understand and consequently calibrate the detector response. The multilevel energy calibration process is outlined in the following with a focus on the derivation of the level 3 absolute residual corrections using events in which such a signature of interest is paired with a Z boson, decaying into a pair of electrons or muons.

5.1. Jets

The particles in the final state created in a collision at the LHC, which are stable enough to interact with the detector material, are either charged leptons or the hadrons originating from the hadronization process of quarks and gluons, due to the phenomenon called color confinement as introduced in section 3.2. Those hadrons, owing to the momentum of the initiating quark and momentum conservation, travel in the same direction in the detector, forming a collimated stream of particles. This object, leaving a complex detector signature, is called jet. The reconstruction of jets is done using jet clustering algorithms, described below. These jets can then be used to study the properties of the initiating quark.

Depending on the underlying process, jets can be very numerous in the event. With their energies often reaching several hundred GeV, they are most important to get a handle on the kinematic properties of the event. Gaining a thorough understanding of jets in the detector is significant for mostly all analyses.

5.1.1. Jet Clustering Algorithms

Jets are composite objects, which are built from the reconstructed PF candidates as introduced in section 4.4.1. These types of jets are referred to as *reconstruction-level* or *detector-level* jets. With MC simulations, the stable particles at the end of the simulation chain prior to the detector simulation can also be used as an input to the jet clustering algorithms discussed in the following. As these objects are composed of the stable particles, they are called *particle-level* jet and can be studied with great detail up to the constituting particles.

The jet clustering algorithm takes the individual PF candidates or MC particles and combines them to form the jets. The combination depends on the transverse momenta of the constituents p_T^i and their separation to each other in the η-φ-plane. This separation ΔR_{ij} is defined by:

$$\Delta R_{ij}^2 = \left(\eta_i - \eta_j\right)^2 + \left(\varphi_i - \varphi_j\right)^2 . \tag{5.1}$$

In the most common sequential recombination jet algorithms [34] the distance between two particles i and j is defined as

$$d_{ij} = \min\left(p_{\mathrm{T},i}^{2a}, p_{\mathrm{T},j}^{2a}\right) \frac{\Delta R_{ij}^2}{R^2}. \tag{5.2}$$

The radius parameter R determines the size of the jets, hence defining the minimal distance of two jets and can be freely chosen.

The clustering is performed sequentially, starting with the pairwise distances d_{ij} of all candidates, as well as the candidates' distances to the beam d_{iB}, which is defined by

$$d_{iB} = p_{\mathrm{T},i}^{2a}. \tag{5.3}$$

Following this metric, the closest candidates ($\hat{\imath}$, $\hat{\jmath}$) are merged into the object k by adding the respective four-vectors. This object is added to the list of candidates, and the constituents $\hat{\imath}$ and $\hat{\jmath}$ are in turn removed. If the minimal distance is $d_{\hat{\imath}B}$, the candidate $\hat{\imath}$ is defined as a jet and also removed from the list of candidates. This process is repeated until all candidates are assigned to jets.

The parameter a is used to decide in which order the candidates are effectively merged. Depending on the sign of a, the algorithm starts by clustering either the softest or hardest candidates in terms of transverse momentum. With $a = 1$, the clustering starts with the softest candidates, which is called the k_T-algorithm. [35] With $a = 0$, the clustering does not depend on the transverse momentum of the candidates but rather only considers their distance in the η-φ-plane. This is the so-called Cambridge/Aachen algorithm. [36, 37]

With $a = -1$, hard particles, i.e. particles with large momenta, are preferred in the ordering of the clustering algorithm. This is the so-called anti-k_T-algorithm. [34]

Jet clustering algorithms should be collinear and infrared safe. Collinear safety means that any of the four-vectors in the jet can be divided into multiple vectors, which add up to the original one, without changing the result of the clustering. An infrared safe algorithm allows

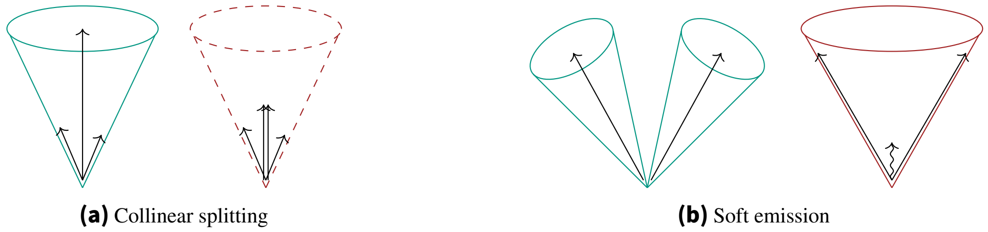

Figure 5.1.: Examples for a collinear (a) and infrared (b) unsafe jet clustering. A collinear-safe jet clustering algorithm is not susceptible to the collinear splitting of the four-vector of any of its constituents. In the example, the jet would not be clustered if the central four-vector is split. Infrared safe algorithms do not change the jets if soft particles are added. Here, because of the added soft particle, the algorithm erroneously merges the two clustered jets. Figure adapted from reference [13].

for an arbitrary number of low energetic particles to be added anywhere to the event without changing the outcome of the clustering algorithm, i.e. the number of jets and their properties do not change by adding "infrared" particles. The above described jet clustering algorithms are called *sequential recombination algorithms* and both are collinear and infrared safe. Examples for collinear and infrared *un*safe jet clusterings are illustrated in fig. 5.1.

The default jet clustering algorithm at the CMS experiment is the anti-k_T-algorithm with the radius parameter $R = 0.4$, which is also used in the later analysis. For certain studies, however, anti-k_T clustering with a radius parameter of $R = 0.8$ is also used.

5.1.2. Pileup Mitigation

To suppress the contribution from pileup (PU), as introduced in section 4.4.2, to the jet energy, CMS has established two different approaches: charged hadron subtraction (CHS) and pileup per particle identification (PUPPI). [38] Both aim at removing particles from PU events from the clustering in order to better describe the event. In terms of PU mitigation, the vertex with the largest sum of p_T^2 of associated objects is called leading vertex (LV). An object is considered "associated" if its track was used in fitting the vertex as introduced in section 4.4.

The most used algorithm to remove in-time PU at the CMS experiment is charged hadron subtraction (CHS). The algorithm associates each charged candidate to a primary vertex and removes the ones originating from PU vertices, i.e. not the LV. CHS helps to drastically reduce the contribution of PU vertices to the jet energy.

CHS, however, is only able to reduce the impact of charged particles contributions. Contributions from neutral particles are alleviated per event by subtracting the average estimated over numerous events. An approach seeking to improve on this is the PUPPI method. It assigns each candidate a weight on how likely it originates from the primary vertex. Tracks from charged candidates are checked if they originate from the primary vertex or close to it. If so, they are assigned a weight of 1, if not, a weight of 0. For neutral candidates, a p_T-weighted distance to all other candidates is calculated and transformed into a probability to originate from the primary vertex.

Pileup Jet Identification

Pileup jets can, although individually rather soft, add up to jets which are reconstructed as objects with substantial p_T. Such jets contaminate the event and need to be removed. By checking each jet's compatibility with the primary interaction vertex of the event, PU jets originating from a single PU vertex can be identified. If a jet, however, contains particles from multiple PU vertices, it has a different shape which can be used to distinguish such jets. Generally, several jet properties are used in multivariate analysis techniques to identify PU jets.

5.2. Missing Transverse Momentum

The partons participating in the collisions at the interaction point have negligible transverse momentum. Because of this, the final state particles are expected to be balanced in the transverse plane:

$$0 = \sum_{\substack{\text{final state}\\ \text{particles } i}} \vec{p}_\mathrm{T}^{\,i}. \tag{5.4}$$

However, as not all particles can be detected and detector noise and PU spoil the measurements of single collision events, eq. (5.4) with any measurement is essentially never fulfilled. The following discussion loosely follows reference [39].

The imbalance is accounted for with the quantity missing transverse momentum (MET), $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$. A commonly used definition is

$$\vec{p}_\mathrm{T}^{\,\mathrm{miss,\,raw}} = -\sum_{i\in\mathrm{all}} \vec{p}_\mathrm{T}^{\,i}, \tag{5.5}$$

which is called *raw MET* and sums over all particles. This, however, includes particles not directly measurable, i.e. "true" imbalances in the transverse plane, but also contributions due to detector noise, reconstruction inefficiencies, faulty detector components, and miscalibrations. To correct the raw MET to converge to the true MET, CMS has several levels of correction, which are applied on top of the raw MET and filters to sift badly reconstructed events, as described in the following.

5.2.1. MET Corrections

Besides the raw MET, eq. (5.5), other types of METs are defined for user analysis. First, the set of particles that is summed over is split into reconstructed jets and unclustered particles

$$\vec{p}_\mathrm{T}^{\,\mathrm{miss,\,raw}} = -\sum_{i\in\mathrm{jets}} \vec{p}_\mathrm{T}^{\,i} - \sum_{i\in\mathrm{uncl.}} \vec{p}_\mathrm{T}^{\,i}. \tag{5.6}$$

Next, replacing the raw reconstructed jets with the corrected jet collection (denoted by the superscript JEC) yields the so-called *type-I corrected MET*

$$\vec{p}_{\mathrm{T}}^{\,\text{miss, type-I}} = -\sum_{i\in\text{jets}} \vec{p}_{\mathrm{T}}^{\,i,\,\text{JEC}} - \sum_{i\in\text{uncl.}} \vec{p}_{\mathrm{T}}^{\,i} \tag{5.7}$$

$$= \vec{p}_{\mathrm{T}}^{\,\text{miss, raw}} - \sum_{i\in\text{jets}} \left(\vec{p}_{\mathrm{T}}^{\,i,\,\text{JEC}} - \vec{p}_{\mathrm{T}}^{\,i,\,\text{raw}}\right). \tag{5.8}$$

Another type of MET is obtained by replacing the reconstructed objects with CHS corrected objects (denoted by the superscript CHS) in the definition of the raw MET. This gives the *type-0 MET*

$$\vec{p}_{\mathrm{T}}^{\,\text{miss, type-0}} = -\sum_{i\in\text{all}} \vec{p}_{\mathrm{T}}^{\,i,\,\text{CHS}}. \tag{5.9}$$

As stated above, this is not necessarily a good correction of the MET, as it considers only charged particle contributions. For a more thorough description, it is assumed that the sum of true transverse momenta for PU equals zero. Furthermore, the energy of neutral PU is presumed to be scaled by a factor R^0. With this, the *type-0 MET* can be written as

$$\vec{p}_{\mathrm{T}}^{\,\text{miss, type-0}} = \vec{p}_{\mathrm{T}}^{\,\text{miss, raw}} + \left(1 - R^0\right) \sum_{i\in\text{ch. PU}} p_{\mathrm{T}}^{i}. \tag{5.10}$$

However, it must be kept in mind that applying both corrections, type-I and type-0, to the raw MET can overcorrect some particle candidates. This is expected to be negligible.

5.2.2. MET Filters

MET is interesting for many physics processes, hence its correct description is of major importance for many analyses. As imbalances can also originate from detector noise, beam halo, or reconstruction deficiencies, those have to be identified. Major effort is put into developing algorithms for the detection of this “fake MET”. For end-user analysis, a set of filters is provided for easy removal of events with such contributions to the MET.

In the later derivation of the level 3 residual corrections, the following filters, described in more detail in reference [39], are applied on recorded and simulated events:

goodVertices This filters events with vertices not compatible with certain quality criteria.

globalSuperTightHalo2016Filter This filter suppresses events with significant contributions from particles from the beam halo, interacting with calorimeter cells.

HBHENoiseFilter and HBHENoiseIsoFilter These filters act on HCAL channels, removing events with anomalous energy deposits. The latter additionally uses ECAL channels to identify activity in HB and HE which does not match tracker measurements.

EcalDeadCellTriggerPrimitiveFilter This filter suppresses events with underestimated energy, due to defective ECAL cells.

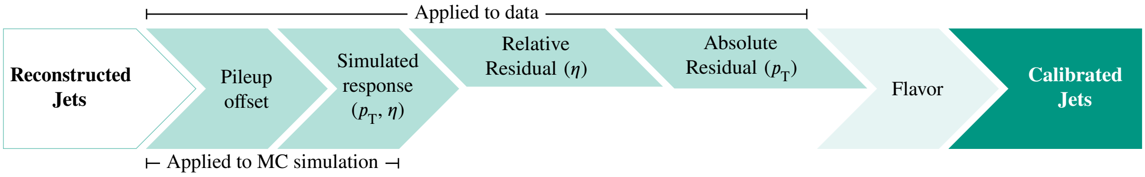

Figure 5.2.: The factorized workflow of the jet energy calibration. Sequential correction factors target different effects. Some corrections are derived in bins of p_T and η as indicated. Residual corrections are applied only on data. The figure is adapted from [40].

BadPFMuonFilter This filters events with muons of low reconstruction quality and high transverse momentum.

5.3. Jet Energy Calibration

Starting from PF particles, jets are clustered using the jet clustering algorithms discussed above in section 5.1.1. As such, jets are very sensitive to measurements of the energy of the comprising particles but also to effects like pileup and the underlying event. The energy of the jet has to be corrected for those effects, as discussed in the following. This workflow is called the jet energy calibration (JEC).

The correction of the jet energy at the CMS experiment is done in a factorized approach. First, using data and simulated events, corrections for the contribution of pileup are derived. Then, corrections of the detector response are determined. Those are referred to as *level 1* (L1) and *level 2/level 3* (L2L3) corrections, respectively. Residual differences are mitigated in the steps afterwards with data-driven methods using dijet (*level 2 residuals*, L2Res) as well as γ + jet, Z + jet, and multi-jet events and the two jets from a hadronically decaying W boson from semileptonic $t\bar{t}$ events (*level 3 residuals*, L3Res). The complete workflow from reconstructed, uncalibrated jets to energy corrected, calibrated jets, ready for user analysis, is shown in fig. 5.2. Indicated in the figure is the optional step in which flavor-dependent corrections are applied on the jets. These corrections are not derived in the jet energy calibration in the current version.

This factorized approach can be written as

$$\vec{p}^{\,\mathrm{corr}} = C \cdot \vec{p}^{\,\mathrm{raw}}, \tag{5.11}$$

with the correction factor

$$C = C_{\mathrm{L1}}(\vec{p}_{\mathrm{T}}^{\,\mathrm{raw}}) \cdot C_{\mathrm{L2L3}}(\vec{p}_{\mathrm{T}}^{\,\prime}, \eta) \cdot C_{\mathrm{L2Res}}(\eta) \cdot C_{\mathrm{L3Res}}(\vec{p}_{\mathrm{T}}^{\,\prime\prime}). \tag{5.12}$$

Denoted like this, the transverse momenta used for derivation of the different correction levels are corrected with up to the respective correction level, that means $\vec{p}_{\mathrm{T}}^{\,\prime} = C_{\mathrm{L1}} \cdot \vec{p}_{\mathrm{T}}^{\,\mathrm{raw}}$ and $\vec{p}_{\mathrm{T}}^{\,\prime\prime} = C_{\mathrm{L2L3}} \cdot C_{\mathrm{L2Res}} \cdot \vec{p}_{\mathrm{T}}^{\,\prime}$.

5.3.1. Pileup Offset Corrections

As described in section 4.4.2, traces from other collisions of the same bunch crossing (in-time), as well as contributions due to a finite reaction time of the detector and the readout from collisions from previous and subsequent bunch crossings (out-of-time), can contaminate the event. The first way of mitigating such pileup contributions is the CHS method, as described in section 5.1.2. Remaining contributions to the jet energy are called *pileup offset*. The level 1 pileup offset corrections aim at reducing this effect. This is accomplished through the *hybrid jet area* and *random cone* methods.

The correction factor of the hybrid jet area method is defined as

$$C_{\text{PU, hybrid}}\left(p_{\text{T}}^{\text{raw}}, \eta, A, \rho\right) = 1 - \frac{\left[\rho_0(\eta) + \rho\beta(\eta)\left(1 + \gamma(\eta)\log\left(p_{\text{T}}^{\text{raw}}\right)\right)\right] \cdot A}{p_{\text{T}}^{\text{raw}}}, \tag{5.13}$$

with the jet area, A, its transverse momentum, p_{T}, and its pseudorapidity ,η. The offset energy, ρ, is determined per event by comparing multi-jet samples, simulated with and without PU, as the average p_{T}-difference of the jets in bins of η and φ. The parameters $\rho_0(\eta)$, $\beta(\eta)$, and $\gamma(\eta)$ relate the correction to the pseudorapity.

Remaining offsets in data are mitigated with the random cone method. For this, a zero-bias sample is used, which consists of events triggered randomly. In the events, jets are clustered in cones placed randomly in the $\eta - \varphi$-plane. The average transverse momentum of these jets is thus assumed to be mainly from PU, disregarding noise. The obtained distribution of the average transverse momentum can then be compared to simulation and the resulting scale factor used to correct data.

5.3.2. Simulated Response Corrections

The second step in the factorized approach of calibrating the jet energy corrects for mis-modelings in the MC simulation of the detector response. For this, reconstructed jets are matched to the initiating generator level particle using the ΔR-metric in the η-φ-plane defined in eq. (5.1). The response is defined as the ratio of the transverse momentum of the reconstructed jet and the transverse momentum of the corresponding particle on generator level:

$$R_{\text{L2L3}} = \frac{p_{\text{T}}^{\text{reco}}}{p_{\text{T}}^{\text{gen}}}. \tag{5.14}$$

It is derived in bins of particle level transverse momentum $p_{\text{T}}^{\text{gen}}$ and η. The correction factors are obtained per bin by averaging the responses within those bins:

$$C_{\text{L2L3}} = \frac{1}{\langle R_{\text{L2L3}}\rangle\left[\eta, p_{\text{T}}^{\text{gen}}\right]}. \tag{5.15}$$

5.3.3. Residual Corrections

The preceding levels of the JEC assume a good description of the detector and its response in the generation of the MC simulation samples. Residual differences due to shortcomings in the MC simulation samples, are hence alleviated using data-driven methods. For this, well-known physics observables like masses are used for calibration. Responses are defined in two different ways, which are explained in the following.

p_T-balance A straightforward approach to balance the probe jet against a reference object that can be measured precisely is called p_T*-balance* or *direct balance*. Assuming no additional objects, the response is determined by measuring the transverse momenta of two objects in a back-to-back topology. The response is then defined by the ratio of the probe and the reference object:

$$R_{p_\mathrm{T}^\mathrm{bal}} = \frac{p_\mathrm{T}^\mathrm{probe}}{p_\mathrm{T}^\mathrm{ref}}. \tag{5.16}$$

The simplicity of this, however, has the obvious drawback of being prone to the existence of imbalances in the transverse plane. Missing transverse energy for example can introduce large biases in the responses. To get a handle on this and to be able to extrapolate the corrections for ideal event topologies, a measure on additional jet activity is introduced as

$$\alpha = \frac{p_\mathrm{T}^\mathrm{jet3}}{p_\mathrm{T}^\mathrm{avg}}, \tag{5.17}$$

where $p_\mathrm{T}^\mathrm{jet3}$ is the transverse momentum of the third-hardest jet and $p_\mathrm{T}^\mathrm{avg}$ is the average momentum of the two leading jets. By selecting only events with a value of α lower than a certain threshold, events with a close to ideal event topology are used for analysis.

MPF A different approach is the missing transverse energy projection fraction (MPF). For this, again, an ideal back-to-back event topology on generator level is assumed:

$$\vec{p}_\mathrm{T}^\text{gen, ref} + \vec{p}_\mathrm{T}^\text{gen, probe} = 0. \tag{5.18}$$

However, when measuring the objects, the responses are scaled by detector response factors, R_ref and R_probe. These non-ideal responses manifest in a missing transverse energy component

$$0 = \vec{p}_\mathrm{T}^\mathrm{miss} + R_\mathrm{ref} \cdot \vec{p}_\mathrm{T}^\text{gen, ref} + R_\mathrm{probe} \cdot \vec{p}_\mathrm{T}^\text{gen, probe}. \tag{5.19}$$

With the assumption of accurately known reference objects, i.e. $R_\mathrm{ref} = 1$ and $\vec{p}_\mathrm{T}^\text{gen, ref} = \vec{p}_\mathrm{T}^\mathrm{ref}$ as well as eq. (5.18), eq. (5.19) reads

$$0 = \vec{p}_\mathrm{T}^\mathrm{miss} + \vec{p}_\mathrm{T}^\mathrm{ref} - R_\mathrm{probe} \cdot \vec{p}_\mathrm{T}^\mathrm{ref},$$

from which the response factor $R_{\mathrm{MPF}} = R_{\mathrm{probe}}$ can be derived as

$$R_{\mathrm{MPF}} = 1 + \frac{\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} \cdot \vec{p}_{\mathrm{T}}^{\,\mathrm{ref}}}{\left(\vec{p}_{\mathrm{T}}^{\,\mathrm{ref}}\right)^2}. \tag{5.20}$$

With this response definition, additional jet— or in this case also called second jet— activity is defined by

$$\alpha = \frac{p_{\mathrm{T}}^{\mathrm{jet2}}}{p_{\mathrm{T}}^{\mathrm{ref}}}. \tag{5.21}$$

Relative Residual Corrections

Relative corrections on the jet energy are derived using jets from the central region, i.e. $|\eta| < 1.3$, where the detector is assumed to be more accurate in measuring the energies of jets. Using the p_{T}-balance method on dijet events, a jet at arbitrary η is balanced against a reference jet within the central region. The correction factors are obtained in bins of η, comparing responses in data and MC simulation.

Absolute Residual Corrections

On top of the relative corrections, absolute corrections are applied. These level 3 absolute residual corrections determine the absolute jet energy scale. For this, jets are balanced against other well known and precisely measurable objects within the central region. As described above, at CMS, γ + jet, $\mathrm{Z}(\rightarrow ee/\mu\mu)$ + jet, and multi-jet events as well as the two jets from a hadronically decaying W boson from semileptonic $t\bar{t}$ events are used to derive p_{T}-balance and MPF responses. Both of the responses rely on the balance of the objects in the transverse plane, with MPF including the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. To account for these event topologies, the responses are determined in bins of additional jet activity α so as to extrapolate the corrections to no additional activity in the event: $\alpha \rightarrow 0$.

5.4. Absolute Residual Corrections with Z + Jet Events

This section discusses the derivation of the absolute residual corrections as just introduced using events with a jet in association with a Z boson. In this analysis, the Z boson decays in a pair of electrons (ee) or a pair of muons ($\mu\mu$). The calibration is performed with data taken in 2018 during Run II of the LHC. The (ultra-)legacy re-reconstruction (UL) with updated detector calibrations and simulations was used.

5.4.1. Event Samples and Selection

The data samples used for the calibration in this analysis were recorded at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ with the CMS detector. The integrated luminosity of the data samples is $59.83\,\text{fb}^{-1}$. An overview of the datasets is given in table A.1 in the appendix.

The samples in year 2018 were recorded in four different eras, which are referred to as “A”, “B”, “C”, and “D”. Within an era, no significant changes in the experimental conditions are expected, and the data recorded in a single era should be able to be treated equally.

Data samples at CMS are provided in *primary* datasets. This is a collection of events, passing one of certain trigger selections, hence enriched with specific types of events. [41] The datasets selected for this analysis are chosen to contain events with the relevant final state particles, especially a pair of muons, *DoubleMuon*, or a pair of electrons or photons. For 2018, a *DoubleEG* dataset, the corresponding trigger path for a pair of electrons or photons, was not available. Therefore, the more inclusive *EGamma* dataset was used. It includes at least a single electron or photon and will be skimmed for events containing a pair of electrons or photons later in the analysis.

For comparison to the recorded data, simulated events pass through the analysis the same way as recorded events do. The simulated MC sample, given in table A.2 in the appendix, is generated using PYTHIA8 [30] and MADGRAPH [42] as an event generator with the CP5-tune [43]. Pileup events are generated as additional events from an estimated underlying distribution and added to the event contents.

Pileup Reweighting

When generating the MC simulations, for each event, the number of pileup interactions, μ, is drawn from a distribution that is expected in data. At CMS, this is done with the minimum-bias cross-section of $\sigma_{\text{min bias}} = 69.2\,\text{mb}$. However, the true number of pileup interactions cannot be known before data taking. To account for the differences between the simulated and the measured pileup distributions, simulated MC events are assigned a pileup-weight. The weights are derived in bins of μ, reweighting the number of simulated events $N_{\text{MC}}^{\mu_k}$ in bin k to the number of events observed in data, $N_{\text{data}}^{\mu_k}$, in the same bin k, yielding

$$w_{\text{k,PU}} = \frac{N_{\text{data}}^{\mu_k}}{N_{\text{data}}} \cdot \left(\frac{N_{\text{MC}}^{\mu_k}}{N_{\text{MC}}} \right)^{-1} . \tag{5.22}$$

Jet Identification

Jets are clustered from the PF candidates, using the algorithms discussed in section 5.1.1. The jet ID used at CMS aims to identify valid jets while suppressing fake ones. A fake jet is an object that is not a jet, but leaves detector signatures comparable to a jet and is therefore falsely reconstructed as one.

For the jet ID, several selections are performed on the constituting objects. Fractions of the total energy of the jet of different particle types are determined and have to fulfil the criteria given in table 5.1. Because of the different regions in the detector, the jet ID is binned in pseudorapidity-regions of the jet. For some of the regions criteria on the number of constituents of the jet, the charged multiplicity, or the number of neutral particles are also given.

Muon and Electron/Photon Corrections

The derivation of the level 3 residuals requires well described electrons and muons. Hence, corrections are applied before using them for the calibration of the jet energy scale. Muons are corrected for detector misalignments, biases in the reconstruction algorithms, and uncertainties in the magnetic field using the so-called "Rochester" corrections. [44] These are provided centrally by the corresponding muon working group [45].

For electrons and photons, correction factors are also provided to correct for differences between MC simulations and data. For the particles, the energy is scaled to match the energy in the MC simulation, while the resolution in the MC simulation is smeared in order to match the resolution in data. A more detailed discussion can be found in [46]. Again, these corrections are provided centrally by the respective E/γ working group [47].

Event Selection

The CMS collaboration provides a file, commonly called "golden JSON", that lists the events certified by the experiment for use in data analysis. The file excludes periods of data recorded when the quality might be compromised because the accelerator, detector, or DAQ were not working properly. For the legacy campaigns of the year 2018, the valid events correspond to an integrated luminosity of 59.83 fb^{-1}.

For the analysis at hand, events have to be selected which fulfill the very specific topology required for the calibration task. The objects needed for determining the level 3 absolute

Table 5.1.: The CMS jet identification requirements for anti-k_T reconstructed CHS jets with radius parameter $R = 0.4$ (AK4CHS) for the data taking period of 2018.

	$\|\eta\| \leq 2.4$	$2.4 < \|\eta\| \leq 2.7$	$2.7 < \|\eta\| \leq 3.0$	$3.0 < \|\eta\| \leq 5.0$
Neutral hadron fraction	< 0.90	< 0.90	–	> 0.2
Charged hadron fraction	> 0	–	–	–
Neutral EM fraction	< 0.90	< 0.99	> 0.01 and < 0.99	< 0.9
Charged EM fraction	< 0.80	< 0.80	–	–
Muon fraction	< 0.80	< 0.80	–	–
Number of constituents	> 1	–	–	–
Charged multiplicity	> 0	–	–	–
Number of neutral particles	–	–	> 1	> 10

residual corrections are a jet and the reference object, in case of this analysis, a Z boson. The latter is reconstructed from two leptons, ee and $\mu\mu$, hence a careful selection for these objects is also necessary.

The event is expected to have at least one jet, $N_{\mathrm{jet}} \geq 1$, and the leading jet has to pass the jet ID requirements, as given in table 5.1. The jets used for the determination of the level 3 absolute residual corrections are corrected with all the correction levels preceding the one to derive.

The Z boson and the lepton pair it is reconstructed from have to fulfill several criteria to be selected for analysis:

- In the muon channel, events have to pass the di-muon trigger path `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8`, requiring one muon with transverse momentum $p_{\mathrm{T}}^{\mu_1} > 17\,\mathrm{GeV}$ and a second muon with transverse momentum $p_{\mathrm{T}}^{\mu_2} > 8\,\mathrm{GeV}$. Their invariant mass has to be greater than 8 GeV. Additionally, muons have to pass the muon identification and PF isolation as described in [48] at the tight working point. In the offline selection after the full reconstruction, the two muons have to pass $p_{\mathrm{T}}^{\mu} > 20\,\mathrm{GeV}$ yielding high trigger efficiencies. Furthermore, only events where the two leading muons are in the tracker region ($|\eta^{\mu}| < 2.3$) are considered.
- In the electron channel, the selection is analogous: Events have to pass the di-electron trigger path `HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL`, requiring a transverse momentum of $p_{\mathrm{T}}^{e_1} > 23\,\mathrm{GeV}$ ($p_{\mathrm{T}}^{e_2} > 12\,\mathrm{GeV}$) for the leading (sub-leading) electron. With the same reasoning as above, the two leading electrons in the offline selection have to fulfill $p_{\mathrm{T}}^{e} > 25\,\mathrm{GeV}$. Both electrons have to pass the cut-based electron identification at the tight working point as described in reference [49]. In this channel, only events with the electrons in the pseudorapidity range $|\eta^{e}| < 2.4$ are considered.
- In both channels, selection requirements apply on the Z boson, as reconstructed from the two leptons: To be able to calibrate in a well-known region, the mass of the reconstructed Z boson, m_Z, must not deviate too much from the literature value, m_Z^{PDG}: $\left|m_Z - m_Z^{\mathrm{PDG}}\right| < 20\,\mathrm{GeV}$. Furthermore, it is required to have a transverse momentum of $p_{\mathrm{T}}^{Z} > 15\,\mathrm{GeV}$. With respect to the leading jet in the event, the Z boson also has to fulfill $\left|\left|\varphi_Z - \varphi_{\mathrm{jet1}}\right| - \pi\right| < 0.44$. This so-called *back-to-back*-cut ensures that the Z boson indeed balances the leading jet. Also, the additional jet activity α measured by the transverse momentum of the sub-leading jet should not exceed the transverse momentum of the Z boson: $p_{\mathrm{T}}^{\mathrm{jet2}}/p_{\mathrm{T}}^{Z} < 1$.

5.4.2. Input Distributions

For the correct derivation of the level 3 absolute residual correction responses, a validation of the input distributions is important. In the following, some variables of interest characterizing the leading jet, the Z boson, and the missing transverse momentum are presented, each in the di-electron and the di-muon channels. In all cases only shape comparisons are performed,

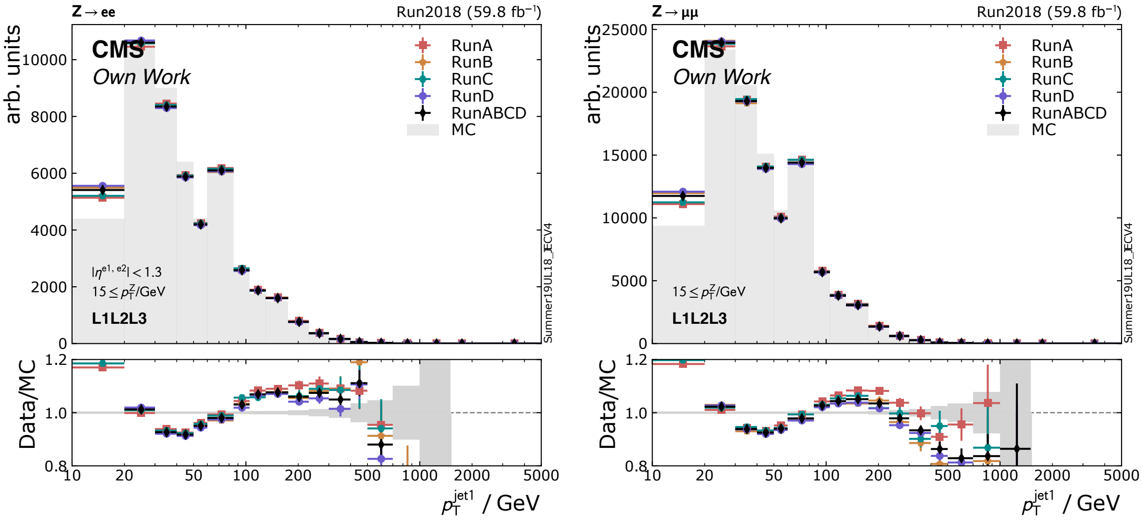

Figure 5.3.: Distributions of the transverse momentum of the leading jet, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

as absolute event counts are not relevant here. Distributions in the simulated MC sample are normalized to the integral of the recorded data for this study. All distributions are shown after the application of the corrections and the selection criteria given above.

Leading Jet

Figures 5.3 and 5.4 show the distributions of the transverse momentum p_T and the pseudorapidity η of the leading jet, respectively. The transverse momentum shows an overabundance of recorded events at lowest p_T. This might indicate an underestimation of PU, resulting in the modulation of the distribution, when normalizing to data. The individual eras behave in a similar way.

The pseudorapidity is reasonably well described. The modulation will be absorbed by the global fit, discussed later, when deriving the global JEC factors in bins of η.

Z Boson

As in the derivation of both responses, MPF and p_T-balance, the Z boson is used to calibrate the leading jet, it is important for it to be well reconstructed. In figs. 5.5 and 5.6 the distributions of the transverse momentum and the invariant mass of the reconstructed Z boson are shown.

As with the leading jet p_T, the distribution of the transverse momentum of the Z boson shows a modulation. A shift in the mass scale of the reconstructed Z boson can be observed in both channels and all runs. This indicates that corrections on the energy scales of the leptons

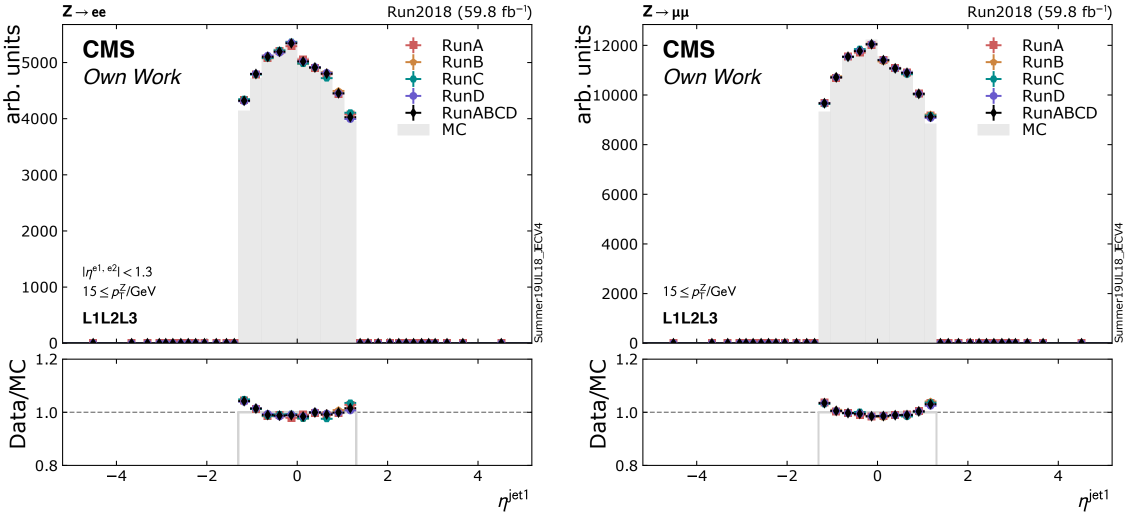

Figure 5.4.: Distributions of the pseudorapidity of the leading jet, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

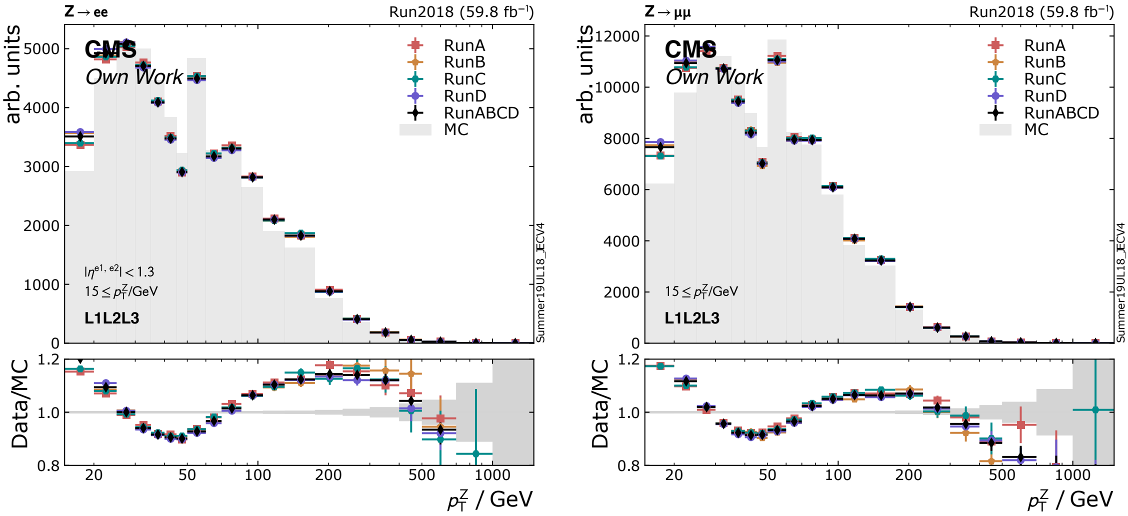

Figure 5.5.: Distributions of the transverse momentum of the reconstructed Z boson, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

used to reconstruct the Z boson are underestimated. This shift in the energy scale results in slightly distorted level 3 jet energy calibrations. Both effects will be mitigated in the final global fit when combining the Z + jet channel with all other analyses providing inputs to the jet energy calibration. Because of this, small deviations are of no major concern at this point.

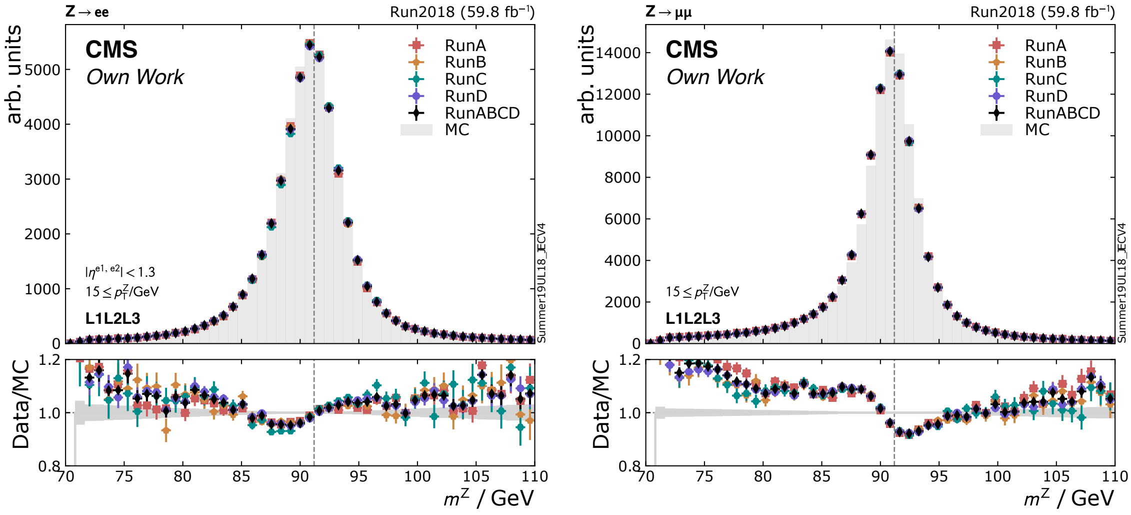

Figure 5.6.: Distributions of the mass of the reconstructed Z boson, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

Missing Transverse Momentum

The distributions of the missing transverse momentum are shown in fig. 5.7. The MPF method projects the reference object's transverse momentum, in this case the Z boson, onto the missing transverse momentum of the event to deduce the responses. Therefore, it is favorable for the missing transverse momentum to be well described. The shape of the simulated samples indeed agrees very well with what was recorded in data. A tendency to overestimate events with low missing transverse momentum, however, can be observed especially in runs A and C in both channels, which may be owed to contributions from PU events not being completely mitigated, hence an underestimation of PU.

5.5. Derived MPF and p_T-Balance Responses

Using the data taken in 2018, corrected, selected, and validated as discussed above, the responses of the p_T-balance and the MPF methods are derived.

In fig. 5.8 the MPF responses for the level 3 residual corrections are shown. The responses follow a roughly Gaussian shape, as is expected for detector responses. All individual runs A to D, as well as the combined run ABCD already show very good agreement in the data/MC ratio in both channels. However, there seems to be a slight shift towards smaller response values in data.

In fig. 5.9 the p_T-balance responses for the level 3 residual corrections are shown. The responses are comparable for all runs, but the agreement with data is not as good as for the

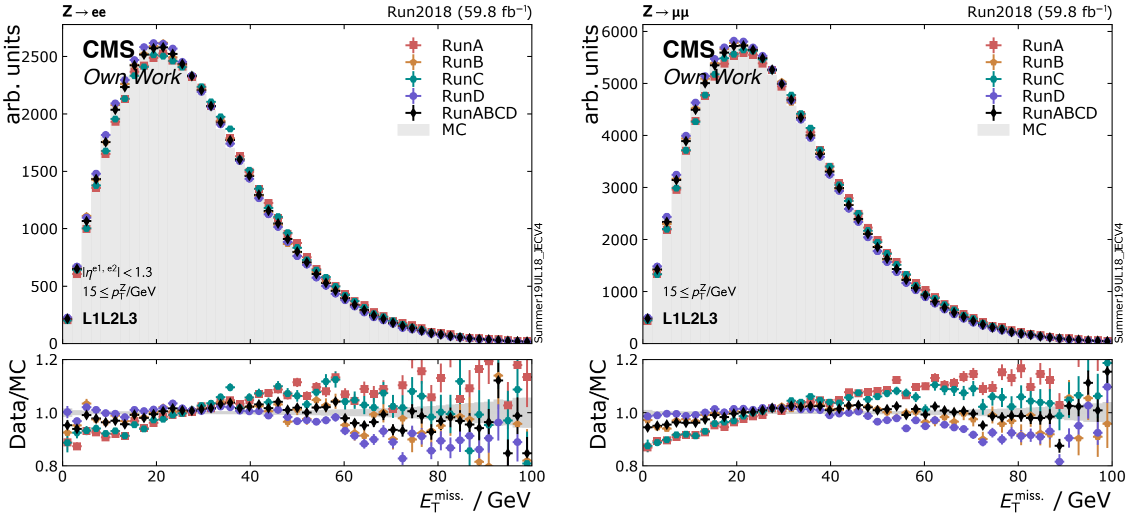

Figure 5.7.: Distributions of the missing transverse momentum in the events, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

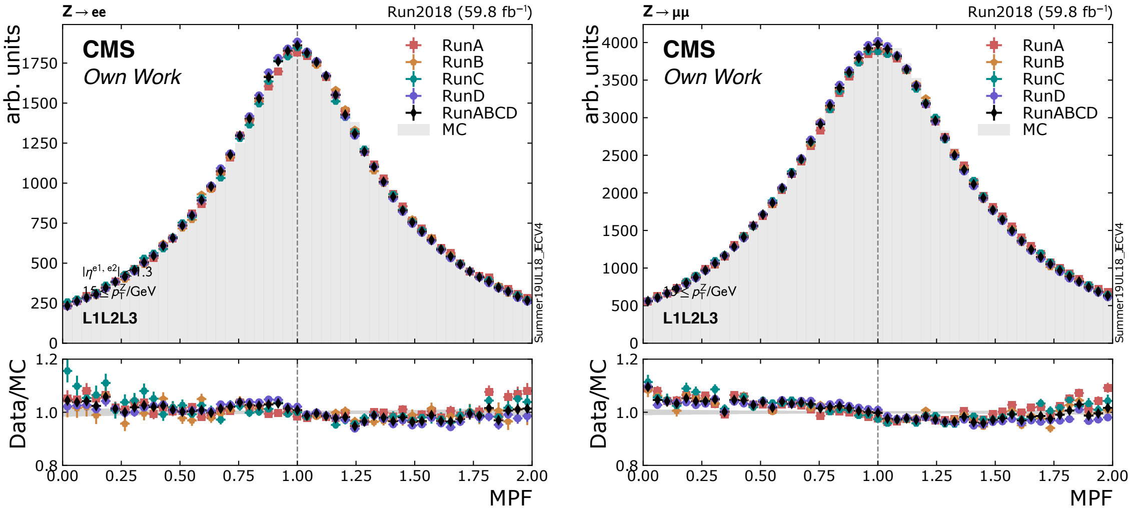

Figure 5.8.: The responses derived with the MPF method on data taken in year 2018, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons. The data-MC agreement is goof.

MPF method. The strong deviations at low values of the response in data is primarily due to Z bosons being wrongly balanced against jets from PU, even though PU jet identification and CHS are applied as introduced above.

Additionally, if the final state hadron, producing the jet in the detector, radiates off a gluon with a significant energy fraction, the resulting particle shower might get reconstructed as a separate jet, hence distorting the energy reconstruction of the final state hadron. This is

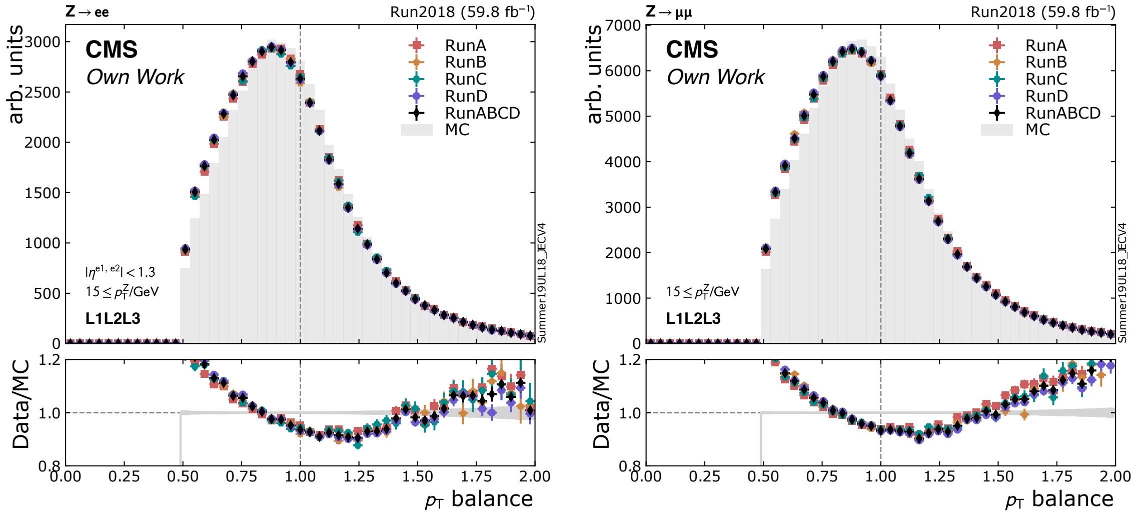

Figure 5.9.: The responses derived with the p_T-balance method on data taken in year 2018, shown on the left in the di-electron, on the right in the di-muon channel. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons. The data-MC agreement is reasonable, considering the MC samples are of leading order only.

called final state radiation (FSR) and causes the response to be shifted to smaller values. Such radiation before the hard interaction is called initial state radiation (ISR).

As can be observed comparing figs. 5.8 and 5.9, the derivation of the MPF is significantly more robust than p_T-balance as it is less sensitive to PU jets and gluon radiation. Because of this, the MPF method is given more weight for the derivation of the correction factors at the CMS experiment.

5.6. Global Fitting Procedure and Validation

The responses, derived with the analysis in bins of the leading jet transverse momentum, p_T^{jet}, and additional jet activity, α, are now used as inputs to a global fitting procedure performed by the *jet energy resolution and correction* (JERC) group of the CMS experiment. The “global fit” combines those responses with the responses derived in the analogous analyses for the absolute level 3 residual corrections using γ+jet and multi-jet events as well as the analysis using the two jets from a hadronically decaying W boson from semileptonic $t\bar{t}$ events. For this, the MPF and p_T-balance responses are extrapolated to $\alpha \to 0$ to describe the events with no additional jet activity. In a χ^2-fit the absolute level 3 residual corrections are determined as a function of p_T^{jet}.

The result of such a global fit is shown in fig. 5.10, taken from reference [50]. This jet response can now be used to correct the absolute jet energy scale for data taken in year

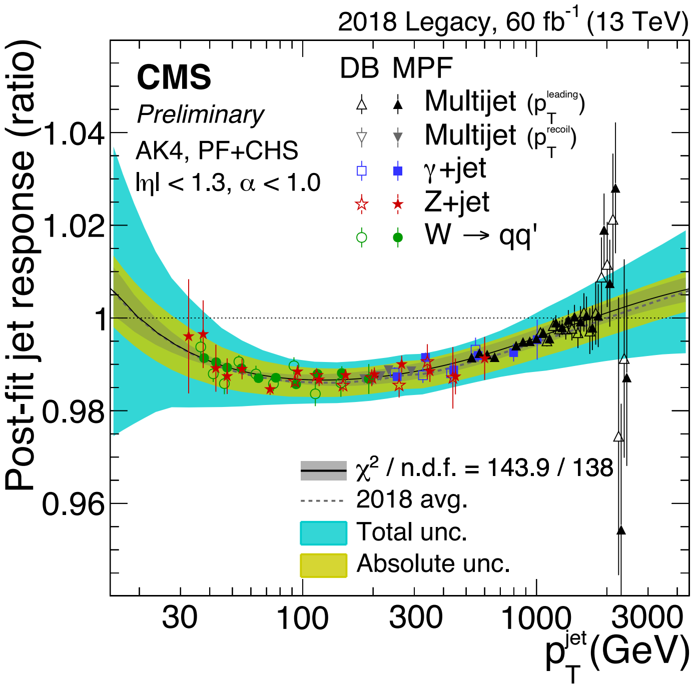

Figure 5.10.: The ratio of the jet response between data and MC simulations after the global fitting procedure for year 2018. The global fit combines the responses derived with the Z + jet, γ+jet, multi-jet, and $W \rightarrow qq'$ analyses, performed by members of the JERC group. Figure is taken from reference [50].

2018. In combination with the relative level 3 residual corrections, they allow for a global correction of the jet energy.

To validate the derived correction factors, in figs. 5.11 and 5.12 their mean values are shown as functions of the transverse momentum of the Z boson and of the absolute pseudorapidity of the leading jet. Each figure contains the distributions without residual corrections applied (top row) and with residual corrections applied (bottom row). It can be observed that the agreement of the data and the MC simulated distributions improves after the application of the residual corrections and is generally within the 2.5 % range. The agreement of the p_T-balance response improves especially for $p_T^Z > 20$ GeV. For smaller values of p_T^Z, the agreement is not as good. The jets using such Z bosons as reference objects are themselves of low transverse momentum. The description of such low energy jets, owing to PU, ISR, and FSR effects, is often deteriorated, causing the slight disagreements for low energies.

In fig. 5.13 the ratios of the mean values of the MPF responses derived in data over the ones derived in MC simulations are shown as a function of the transverse momentum of the Z boson and the pseudorapidity of the leading jet for both channels. Especially in the pseudorapidity region $|\eta^{\text{jet1}}| < 1.3$, the agreement is very good. Only a few bins deviate more than 3 %. This is the detector region used for the derivation of the absolute residual corrections. Outside this region, the agreement is not as good. As discussed above, for low transverse momentum Z bosons, especially in the region at $\eta \approx 1.3$, where they transition from barrel to endcap, the description is very poor.

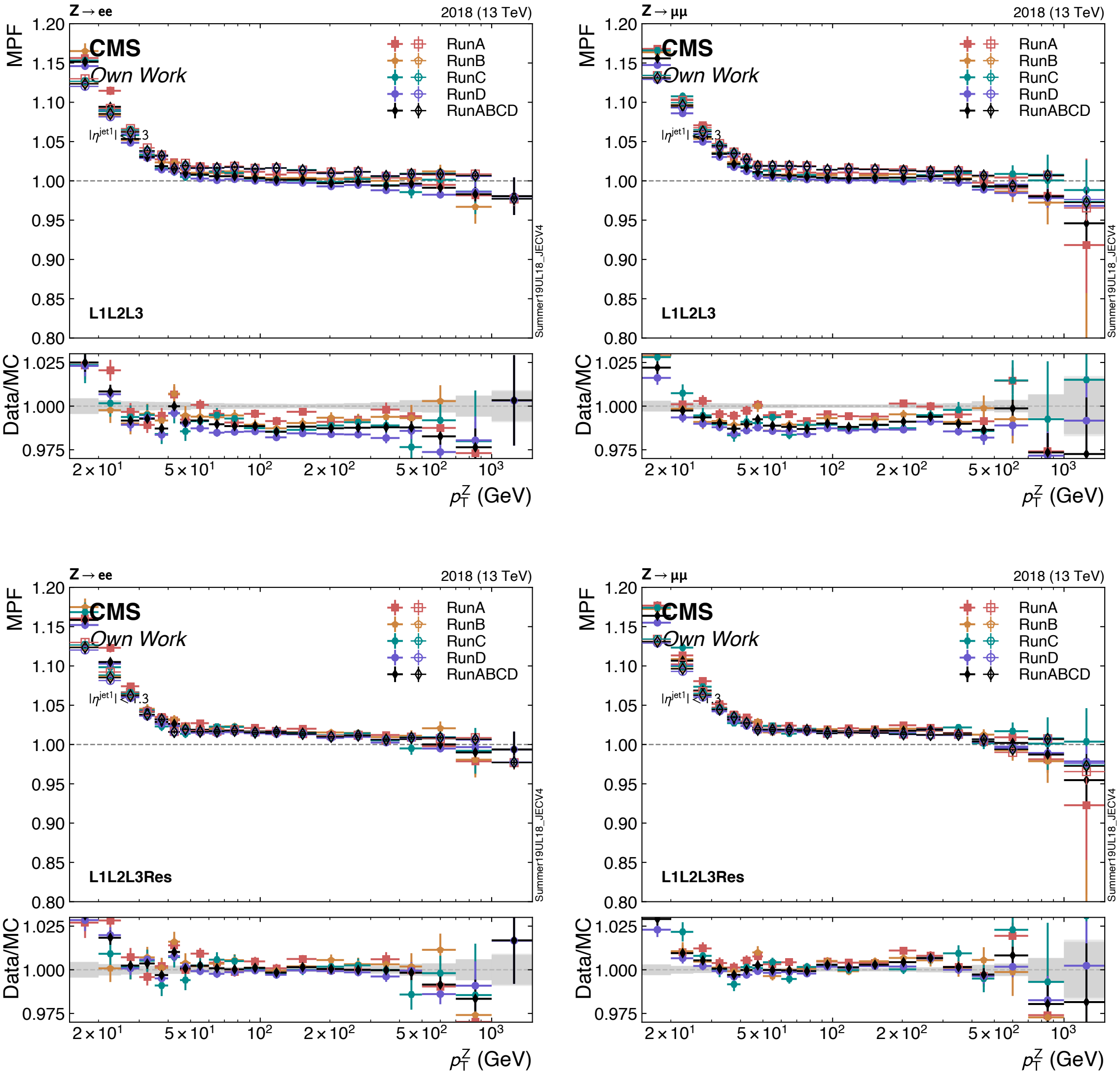

Figure 5.11.: Distributions of the MPF responses as a function of the transverse momentum of the Z boson, shown on the left in the di-electron, on the right in the di-muon channel. The upper and lower plots show the distributions without and with the level 3 residual corrections applied, respectively. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

Figure 5.14 shows the ratios of the mean values of the p_T-balance responses derived in data over the ones derived in MC simulations as a function of the transverse momentum of the Z boson and the pseudorapidity of the leading jet for both channels. The agreement is excellent, especially in the pseudorapidity region $|\eta^{\text{jet1}}| < 1.3$. Only for high p_T^Z in a few bins the deviation is more than 3 %.

Overall, the good agreement of the responses, especially in the regions in which the responses are derived in, validates the absolute level 3 residual corrections and makes them suitable for use in physics analyses. The jet energy calibration resulting from the global fit, determined

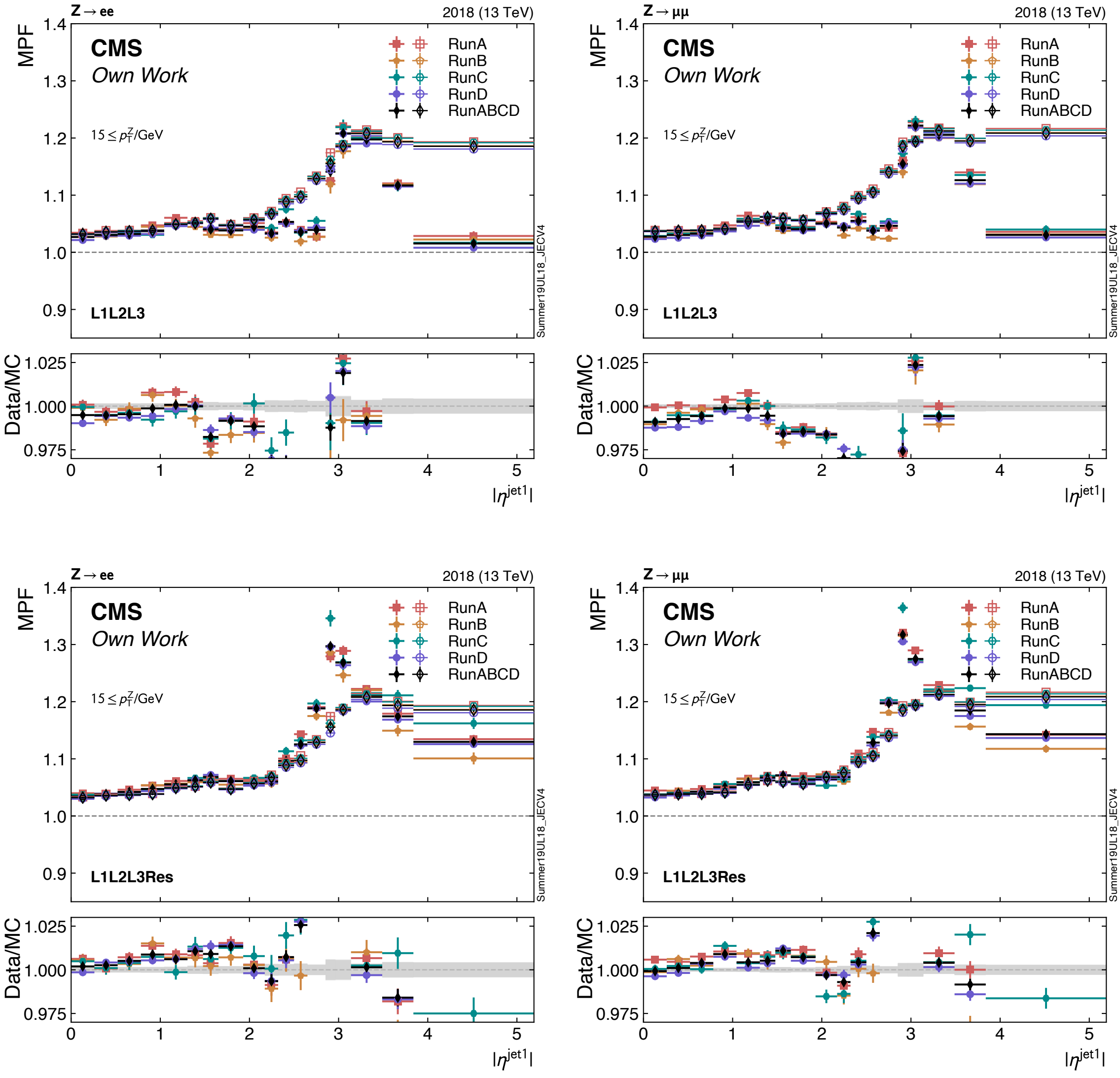

Figure 5.12.: Distributions of the MPF responses as a function of the pseudorapidity of the leading jet, shown on the left in the di-electron, on the right in the di-muon channel. The upper and lower plots show the distributions without and with the level 3 residual corrections applied, respectively. MC distributions are normalized to the integral of the recorded data to allow for shape comparisons.

with the responses derived within this thesis are used in several analyses performed by the CMS collaboration.

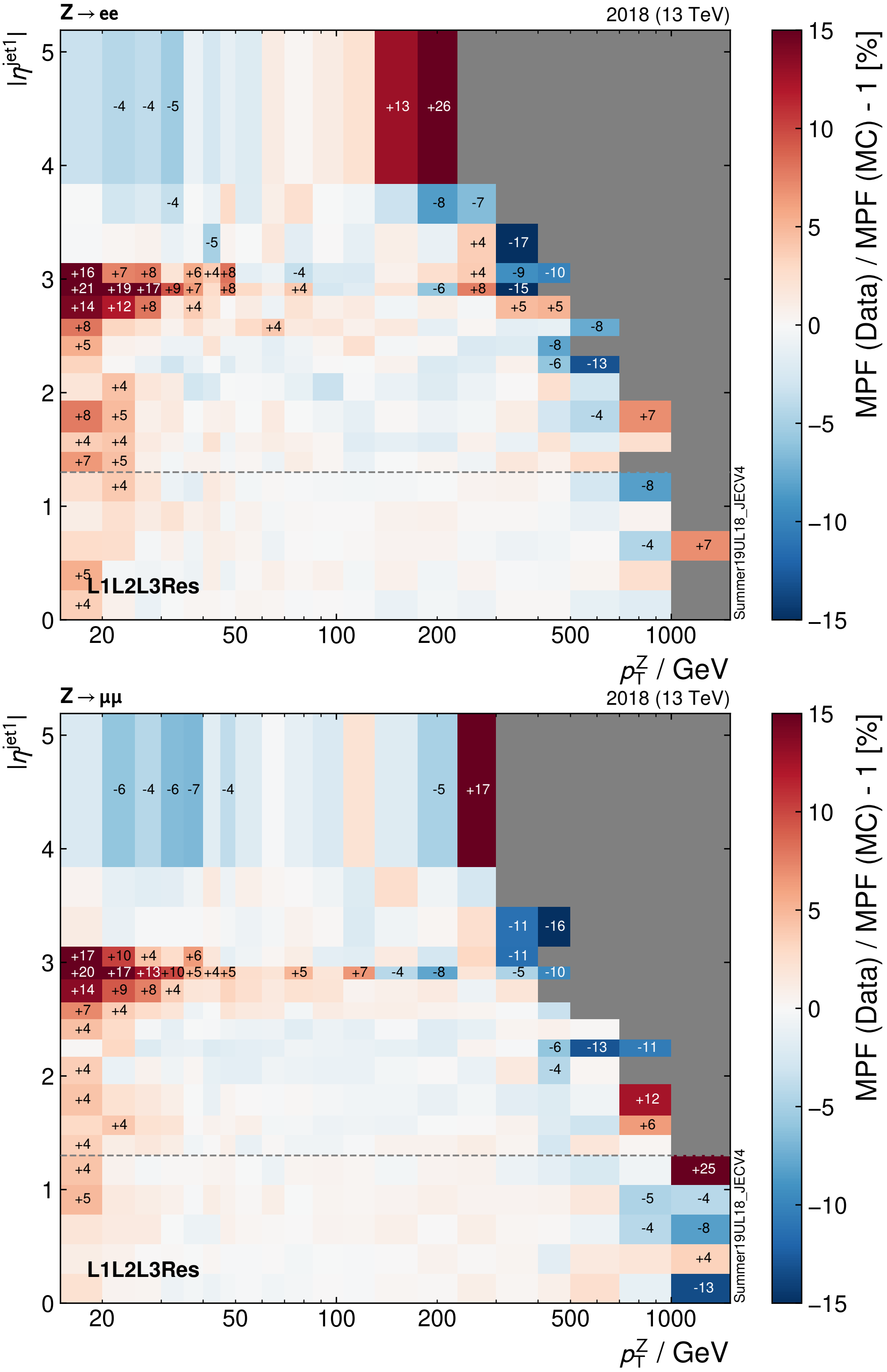

Figure 5.13.: The ratios of the mean values of the MPF response in data over MC as a function of the transverse momentum of the Z boson and the leading jet pseudorapidity, shown on the left in the di-electron, on the right in the di-muon channel. Deviations of more than 3 % are given numerically.

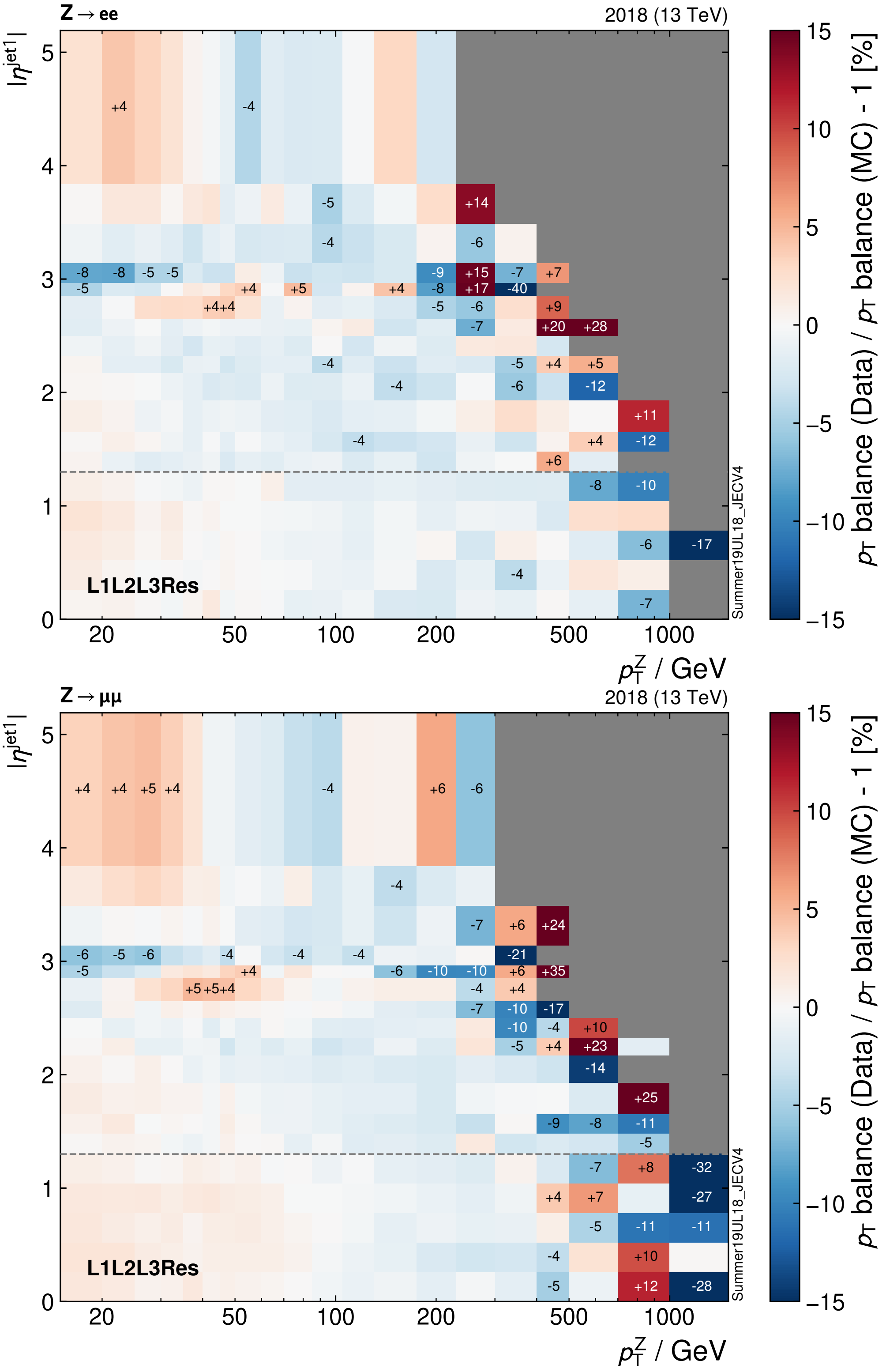

Figure 5.14.: The ratios of the mean values of the p_{T}-balance response in data over MC as a function of the transverse momentum of the Z boson and the leading jet pseudorapidity, shown on the left in the di-electron, on the right in the di-muon channel. Deviations of more than 3 % are given numerically.

6. Resource Scheduling

High energy physics experiments generate a lot of data, which has to be analyzed at different levels: Starting at the detector, at every moment enormous amounts of data are collected. Trigger algorithms, as discussed in section 4.3, perform reconstruction tasks on a compute farm to decide whether to keep an event. The data, passing these selections, is then stored on permanent storage. On a global collaboration of computing centers, the so-called WLCG which was already introduced in section 2.1, the data is then reconstructed, i.e. from the raw detector data, a data format is derived in multiple steps, which is easily usable by physicists to perform statistical data analysis of the billions of events. This process is outlined in section 4.5. Furthermore, the, colloquially called, *grid* is also used for various other computing tasks, such as Monte Carlo simulations, which are shortly discussed in section 4.5.2. This constitutes an enormous need for computing resources. Also the analysis of the simulated events as well as the recorded samples require computing resources. Cross-section measurements, searches for new physics, signal strength measurements, or detector calibration tasks, for example, need to process large datasets, containing thousands of files, summing up to petabytes of data to transfer, handle, and analyze. All this results in very diverse resource requirements for the computing infrastructure, which must be taken into account.

The WLCG, provides a huge amount of such resources for performing analysis, but especially for the two obligations, of official reconstruction, and MC simulation. For these tasks, the experiments have dedicated groups, taking care of the planning, organization, and execution of the campaigns. The groups work together closely with the respective groups responsible for resource planning at the WLCG.

With these liabilities, the CMS experiment expects a shortage of computing time of a factor of three to four in the worst case in 2034, when recording data in Run 5, starting in year 2032 as referenced in fig. 2.2 in chapter 2. Developments in different areas are expected to alleviate this shortage:

1. The improvement of the implemented algorithms, the underlying software, and the computing model.
2. General advances in hardware, especially CPU performance, and the usage of novel hardware architecture like graphics processing units (GPUs).
3. Usage of additional, non-HEP resources.

Item 1 takes place continuously in the HEP community and the groups, implementing the specifics of the tools used for analysis. Working packages were devised, which are addressed in experiment overarching teams already established and working efficiently as of today, e.g. in Germany these developments are done within the project “Erforschung von Universum und Materie” (ErUM) by the *Bundesministerium für Bildung und Forschung* (BMBF), the German federal ministry of education and research. Other countries have their distinct projects for the developments. Item 2 is of general concern, not only in high energy physics, but improvements in CPU performance and the development of special accelerators, as e.g. graphics processing units (GPUs) are of interest for many areas. In this chapter item 3, the provisioning of additional resources, constituting one part of the ErUM project, is discussed.

Research groups are normally associated with research institutes or universities, which provide the groups with the necessary computing resources to their needs for tasks like data analysis, or simulations. This is in addition to resources available to the groups e.g. from their experiment collaborations, as the WLCG in case of HEP. Those sites are usually multipurpose, serving groups from different research areas with diverse needs. More resources might be provided by further parties resulting in a number of sites for the groups to use. With several sites, the user has to decide which site to use for which task at a given time. While each additional site provides computing resources, it also leaves the user with a complex decision to make.

As the sites are operated by different providers, the way how to access and how to use them might differ widely. To gain access depends on the specifics of the policies, implemented at each site. For further discussion on site policies, please refer to section 6.6. Also, each site operates its distinct local resource management system (LRMS) to manage their resources. All this adds another layer of complexity to the efficient usage of multiple of such different sites.

In the end, the user has to decide, which sites to submit a specific workflow to. As each site uses its own LRMS, the user also has to check, whether the site is available at the moment, or whether it is busy running jobs from other users. More than that, by deciding for a specific site and submitting the jobs to the LRMS of that site, the jobs will possibly be sitting in the queue for some time, idling, even though resources might be available for processing jobs at other sites.

To tackle all these challenges, KIT develops a resource management tool, using feedback loops, to dynamically allocate resources at multiple sites whenever needed and to consolidate all these sites within one system, providing a single point of entry to the various sites for the user. The resource management tool, called COBALD/TARDIS [51, 52] integrates all resources into one, common overlay batch system (OBS).

6.1. Batch Systems

Distributing large workflows, consisting of numerous computing payloads, over hundreds of worker nodes, while ensuring fair allocation of resources to different users is task of a *batch system*. Batch systems, well known in the context of high performance, and distributed computing, introduce several concepts to distribute the computing payloads to the worker nodes, managing multiple users: [53]

Job or payload The central concept of most importance is the job, also referred to as payload. It defines the executable and the resources needed for it to run. This includes the number of CPU cores, the amount of memory and possible other resources, like disk storage or accelerators like GPUs necessary to process the job. Often also an upper limit on the expected wall-clock time of the jobs expected processing time has to be provided. A workflow like an analysis, e.g., is most often split into a lot of jobs, performing the same analysis on different input data files.

Queue The batch systems collects the submitted jobs in queues, servicing different user groups and/or different resources. The jobs stay in their queues until they are processed. By checking their queues, a user can also estimate how busy a batch system is.

Worker Nodes A worker node is the abstraction of the physical hardware, representing i.a. CPUs, memory, disk. The machines integrated into the batch system are available for processing the jobs, submitted by the user. Conditions can be implemented, which jobs to accept for processing, as e.g. machines can be assigned only to specific queues, or specific users. By using this abstraction, a machine can be taken out of the batch system and disabled for job processing, without physically disabling it.

Scheduler The scheduler takes care of matching jobs from the queues to machines considering the jobs requirements, machine availabilities, and user priorities aiming at a fair usage of the whole system. Specific policies can be implemented for the scheduling process to follow, if e.g. a specific user should have higher priority.

The collection of all the involved machines, especially the worker nodes is often referred to as *pool*. To increase the overall utilization of their resources while maintaining a fair share of them between users, most sites have a batch system in place for the management of their resources.

HTCondor

A very common batch system in HEP is HTCondor [54]. The HTCondor system operates with several daemons, running on the involved hosts, taking care of the services the hosts should provide. Users define jobs using the job description language (JDL) [55], a key value based description of the job. After submitting the job to the HTCondor instance, several daemons handle the management. The most important ones are

condor_master The `condor_master` is the main daemon, taking care of all other daemons running on the host. When starting HTCondor, it starts all other daemons, the host is configured to run, restarts them if needed, and periodically runs clean up campaigns.

condor_schedd When a user submits a job using the command "`condor_submit`", the job gets registered within the `condor_schedd` running on the submit host. When matching jobs to resources, the set of `condor_schedds` are used to generate the queue.

condor_shadow The condor_shadow is a lightweight process on the submit machine, representing the job. It takes care of the file transfers of the executables to the worker nodes, input files, output files, and redirects standard output and error.

condor_startd The condor_startd is the daemon running on the worker nodes. It starts the jobs using the condor_starter and reports the status of the jobs to the condor_collector.

condor_starter Each job on the worker nodes is taken care of by a condor_starter. It communicates with the condor_shadow on the submitter machine, transfers files, checks the jobs status and reports it to the condor_shadow.

condor_collector The condor_collector obtains and provides information from several HTCONDOR daemons. The condor_schedds and the condor_startds report the status of their jobs and the worker nodes to the condor_collector. The information is then used by the condor_negotiator.

condor_negotiator The matching of jobs to resources is performed by the condor_negotiator. Taking care of the requirements of jobs and machines while also considering user priorities, it decides which job to start on which resource.

6.2. Resource Management with COBALD/TARDIS

The resource management COBALD/TARDIS [51, 52] is developed at KIT to perform the dynamic integration of sites for their transparent usage. The system integrates the resources transparently in an OBS which acts as the single point of entry for the users. TARDIS (Transparent Adaptive Resource Dynamic Integration System) uses the decision logic of COBALD (COBALD – the Opportunistic Balancing Daemon) based on a feedback loop and by this reacts on the decisions of the OBS, as discussed in the following.

Depending on the type of site being integrated, the resources allocation differs: For sites using a LRMS, a placeholder jobs is submitted. Resources at sites using a cloud API for their resource management are allocated by booting a placeholder virtual machine or container. The term LRMS is used in the following in the context of integrated sites to better distinguish from the OBS. A placeholder running on the resource of the integrated site, preparing it for use within the OBS, is coined DRONE. The DRONE is a generalization of the pilot concept as used e.g. with glideinWMS [56] within the WLCG.

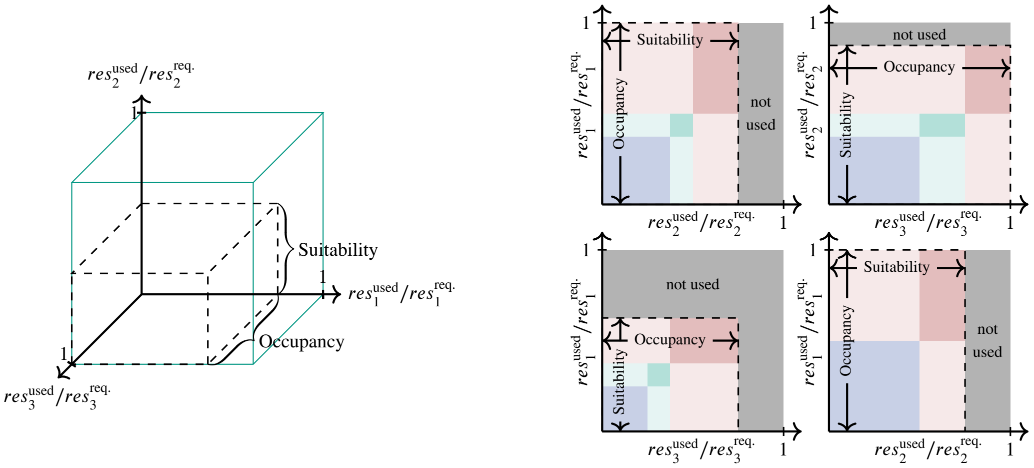
$res_2^{\text{used}}/res_2^{\text{req.}}$
1
Suitability
$res_1^{\text{used}}/res_1^{\text{req.}}$
1
Occupancy
1
$res_3^{\text{used}}/res_3^{\text{req.}}$
Suitability
Occupancy
not used
not used
Occupancy
Suitability
not used
Occupancy
Suitability
Suitability
Occupancy
not used
$res_1^{\text{used}}/res_1^{\text{req.}}$
$res_2^{\text{used}}/res_2^{\text{req.}}$
$res_3^{\text{used}}/res_3^{\text{req.}}$
1

Figure 6.1.: The metrics occupancy and suitability are determined per DRONE. They are a measure for how well the jobs running at the moment fit with the integrated DRONEs. On the left, the determination of the metrics, depicting the ratios $res^{\text{used}}/res^{\text{req.}}$ as a multidimensional volume is shown. The green box shows the volume provided by the DRONE, the dashed box shows the volume used by the jobs. On the right, multiple DRONEs, each with several jobs are shown. The decisions by COBALD are then based on the pool average of the metrics.

6.2.1. COBALD

COBALD assesses the fit of the current job mix to the currently integrated resources by using the metrics

$$\text{occupancy} = \max_{res \in \mathcal{R}} \left(\frac{res^{\text{used}}}{res^{\text{req.}}} \right), \text{ and} \tag{6.1}$$

$$\text{suitability} = \min_{res \in \mathcal{R}} \left(\frac{res^{\text{used}}}{res^{\text{req.}}} \right). \tag{6.2}$$

$\mathcal{R}$ is the set of resources to consider, it normally contains number of CPU cores, and amount of memory, however, other resources as number of GPUs, amount of disk, or any other measurable can be defined. The quantities res^{used} and $res^{\text{req.}}$ correspond to the amount of the various resources used by the jobs, running within the DRONE and requested by the DRONE, respectively. With these definitions, occupancy is a measure for how much of the DRONE is reserved by the jobs from the OBS, as in how much is actually occupied and suitability measures how well the DRONE is utilized, i.e. how well it suits.

In this sense, a DRONE can be perceived as a multidimensional volume in a space spanned by the resources in $\mathcal{R}$. Each job in turn occupies a fraction of this volume, decreasing the remaining inner size of the DRONE. Another job can only fit, if its "shape" fits within the remaining volume. DRONEs which are indistinguishable are grouped in a pool by COBALD and the metrics are averaged per pool. The metrics per drone and per pool are depicted in fig. 6.1. Based on these metrics, COBALD implements the logic when to increase the number of DRONEs at a given site and when to release resources with a poor usage.

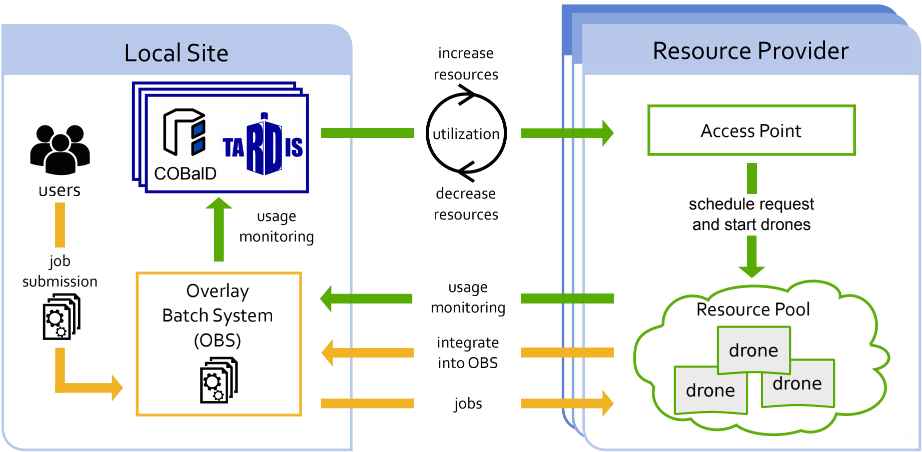

Figure 6.2.: The feedback loop, COBALD/TARDIS uses to schedule resources. The users submit their jobs to the OBS which dispatches them to all available nodes, including the DRONEs running at the different integrated sites. COBALD uses the OBS to assess the metrics occupancy and suitability. Depending on these metrics, the number of resources at the respective resource provider is increased by submitting resource requests to the corresponding access point. The DRONEs starting up, are in turn integrated in the OBS. Figure is taken from [58]

6.2.2. TARDIS

TARDIS performs the integration of the resources into the OBS and takes care of the whole life cycle management of the individual DRONEs. The life cycle is described by the states, a DRONE can be in: from booting state, over running, to stopped, deleted, and error. Through generic functions, DRONEs get deployed, their status is checked, or they are released, as discussed below.

TARDIS allows for a modular implementation of different types of sites and OBSs by providing abstractions for both. For a site to be able to be integrated using COBALD/TARDIS, a `SiteAdapter` for the specific LRMS or cloud API has to be implemented. This is done as a python-class providing asynchronous methods for deploying and stopping a DRONE, getting its status and handling exceptions. The same is done to allow for different OBSs: a python-class implements, how to integrate and disintegrate allocated resources, and how to determine the metrics occupancy and suitability for COBALD to use.

In fig. 6.2 the way COBALD/TARDIS controls the number of resources integrated in the OBS is shown: TARDIS periodically uses the OBS to measure the metrics occupancy and suitability of the resources. In turn COBALD uses these metrics to decide whether to increase or decrease the number of resources at a given site. The user submits their jobs to the OBS as single point of entry, which dispatches them to the individual DRONEs integrated into the OBS as worker nodes. This again has an effect on the metrics occupancy and suitability and in turn affects the decisions made by COBALD/TARDIS. Rather than trying to predict the resource needs as done in the predecessor ROCED [57], COBALD/TARDIS is reacting on the decisions of the OBS, utilizing a feedback loop.

6.2.3. COVID-19

The conceptual design of COBALD/TARDIS enables their use not only for the scheduling of resources in the very specific HEP context, but rather allows to provide resources in a more general meaning. When the COVID-19 pandemic broke out, a major research effort was undertaken to understand the virus and find possible treatments of the disease caused by the virus. For research in this area, the Folding@Home project [59] provides a distributed computing infrastructure to which everyone can provide resources by running a specific software on their own hardware. The Folding@Home software makes the resources available to the distributed computing infrastructure and pulls the computing payloads, submitted by the researchers, from central servers for processing on the local hardware.

With COBALD/TARDIS it was possible to dynamically provide institute resources to the project when not needed otherwise. The setup allowed contributing roughly one hundred CPU cores over the span of a couple of weeks, completing 49 109 work units, a measure of the amount of resources provided by the project. This placed the institute on rank 4686 of 2 908 486 of the whole Folding@Home community. [60]

6.2.4. HTCONDOR Flocking

Another way of using resources from another HTCONDOR pool is to use the HTCONDOR flocking mechanism. This is an HTCONDOR proprietary mechanism, allowing to define a set of other HTCONDOR pools, which jobs can be forwarded to, if no match is found in the local pool. A match is then looked for in any of the other pools, the local pool is configured to flock to. The pools can be set up to flock in both directions. Using this mechanism has the advantage of not needing any further software, however, only works with HTCONDOR pools.

This mechanism was used to make the tier 3 analysis cluster TOPAS, which was designed to be used for high throughput data analysis workflows, available for the users of the institutes batch system. The latter was set up to flock to the HTCONDOR batch system, managing the TOPAS resources. Further details on this setup can be found in references [61] and [62].

6.3. Site Requirements

In order to be able to be used for scientific workflows, sites have to meet very special requirements. They have to provide users with the ability to access the resources in a way, sensible for their workflows to run.

The sites also have to provide users with the ability to access their data for analysis, but also give opportunity to store their results. The requirements to this vary heavily on the size of the respective files. While small datasets summing up to a couple or tens of gigabyte are often able to be stored on the site, processing large datasets in the range of terabytes or even more, requires a different solution. This is either a dedicated large scale storage at the site or the possibility to use network to access external storages. Depending on the other research infrastructure available to the researchers, the latter is normally preferable for an easy integration with already established experiment storage. However, for this to work reliably the network connection to the outside world, especially has to be reasonably sized.

Such access to external resources like storages can give rise to other issues as specific protocols might use non-standard ports. This is also a challenge with OBSs using non-standard ports to connect e.g. the scheduler of the OBS to jobs running on the worker node of the sites. Such ports have to be opened in the firewall of the respective site.

For the dynamic integration of such sites using COBALD/TARDIS, authentication against the site has to be possible in an automated way, i.e. one-time passwords (OTPs) should not be used and multi-factor authentication (MFA) should not be required as this makes human interaction necessary. For sites using such authentication models, different solutions have to be found. Related to this is also the question of accounting on the site in the scope of the multi-user use of the resources. Jobs from the OBS, which get matched to the resources of the site can be of any user, authenticated against the OBS. This user is not necessarily the owner of the account on the site. For accountability and liability reasons, however, it has to be possible to track users who submitted specific jobs. This is discussed in more detail in section 6.6.

Another issue can be the lack of specific software, necessary to perform the analysis and compute jobs. Different ways to provide the missing software exist. They are discussed in greater detail in section 6.4.

Last but not least, the sites should offer a certain amount of resources for the research projects for the work of integrating the site to be justifiable.

6.4. Software Deployment and Virtualization

Many sites not directly associated to an experiment or a specific scientific community lack specific software necessary for running experiment or analysis workflows. As the analysis tools are often very specialized and might need specific versions of software, this has to be provided to the worker nodes on non-community sites. In the WLCG for example it has been agreed upon to use *CentOS Linux* derivatives, which are based on Red Hat Enterprise Linux (RHEL), a Fedora-based [63] distribution as the underlying operating system.[1]

How analysis software can be made available for the jobs depends on the site. If a specific version of operating systems (OSs) is required, this can be distributed as a virtual machine

[1] As of writing this thesis, Red Hat Inc. announced *CentOS Stream* [64], a rolling release distribution, and that it will discontinue the development of *CentOS Linux* [65]. The corresponding working groups in the WLCG have “propose[d] to target CentOS Stream as the standard distribution for experiments” [66] on October, 25 2021.

(VM) image if the site offers some type of cloud infrastructure management like OpenStack [67] or CloudStack [68]. If this is not the case, but the site offers the possibility to use container technology, container images of appropriate architecture can be used. This is described in more detail later in this chapter. Both types of images, VMs or containers, can also be used to ship the necessary software as well.

Another way to provide the software is by making the software available on the site itself. Either by installing the software in the user space, or, especially for software with rolling updates, by using special software repositories, as for example the one widely used in HEP, CVMFS, as described in the next section.

6.4.1. CERN Virtual Machine File System

The CERN virtual machine file system (CVMFS) [69, 70] is a read only file system developed at CERN for distributing software across the world. It is designed in a tiered structure, allowing for aggressive caching on the all levels.

A central stratum 0 web server is used for managing the software which should be published. While data can be accessed by the clients directly from a stratum 0, aiming for more stability stratum 1 servers can be set up for data redundance. Mirroring the stratum 0 data, hence acting as proxies, deployment of stratum 1 servers constitutes a more reliable infrastructure.

On the client machines, the file-system is mounted as a read-only file system. The communication to the strata is done using the standard hypertext transfer protocol (HTTP). This has the advantages of being able to reuse technologies, developed for content distribution within the worldwide web. Especially, it allows to place additional squid-proxies [71] enabling further caching. Also on the client machines, mounting the file-system, caches can be instantiated. In an infrastructure with performant parallel file systems, those can be used for caches as well.

6.4.2. Virtual Machines and Containers

The environment, the software is running in can be provided through virtualization technologies like virtual machines (VMs) or containers. VMs allow for running any guest system in terms of hardware, OS, and software on a host of the same processor architecture, no matter the host's OS. This virtualization is done on top of a hypervisor running on the host, taking care of the communication of the host's and the virtualized hardware. Because of this design, VMs are very versatile and can be run in very different environments, however hardware abstraction and emulating the guest OS can be very costly in terms of computing power.

A more resource-efficient approach is the usage of container technologies. Containers reuse the kernel of the host's OS and only ship the needed libraries and application layers. They are started through a lightweight container engine, which if necessary decompresses the image and prepares required mount points. Containers are also favorable because their images are

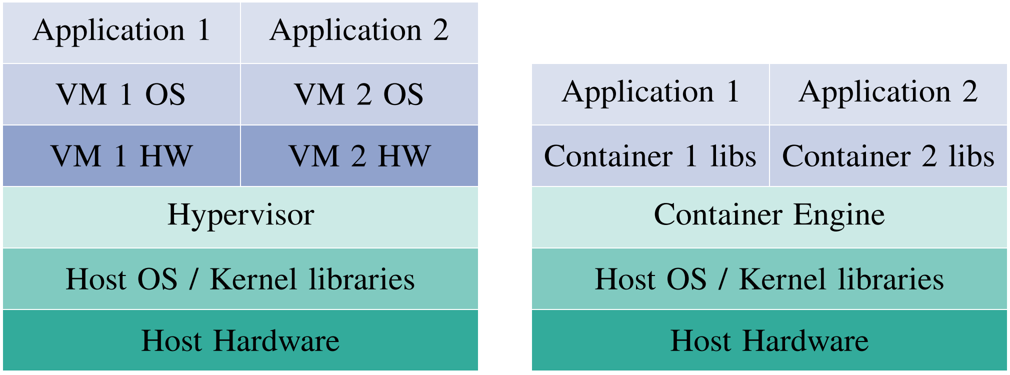

Figure 6.3.: Different virtualization technologies: virtual machines (VMs) on the left, containerization on the right. Not virtualized components are shown in shades of green, virtualized components in shades of blue. On the left, applications run in virtual machines, which are managed by the hypervisor. Depending on the type of hypervisor, no OS is needed on the host (type 1). The layer is shown as the sites discussed here using type 2 hypervisors require an underlying OS. On the right, the applications run in containers. The container images are prepared and started through a lightweight container engine. Figure adapted from reference [72].

a lot smaller, compared to VM images. A comparison of the two technologies is shown in fig. 6.3.

In the scope of this thesis, two different container technologies are considered: DOCKER [73, 74] and SINGULARITY [75, 76]. DOCKER has gained a lot of traction in industry in the last couple of years for deploying software and services: due to its small footprint, a lot of services were "dockerized", meaning they are shipped as DOCKER images and easily set up as containers when needed. This allows to deploy complex infrastructures by combining several of such containers, dynamically starting them up whenever needed and restarting them when something fails. For managing this, tools like KUBERNETES [77] are developed.

SINGULARITY was developed for the science case in the HPC context. The underlying technology is fundamentally different to the one used by DOCKER: The latter uses a root-owned daemon (the container engine) for the container management, however, this implicitly grants root-privilege to every user allowed to run containers which is out of the question for production resources with unknown and untrusted users. For this reason SINGULARITY was developed. SINGULARITY containers can be run in a user's namespace, cf. next section, limiting the process to a certain set of privileges. This is also important in the context of batch systems, where jobs should not be able to use more resources than requested when submitted.

Container images are mostly provided as compressed, layered file systems. Starting from a base image, changes to it are added as thin layers on top, providing the more specific software, or runtime environment. The layered structure allows for different images to reuse common layers, resulting in a small storage footprint. SINGULARITY uses SQUASHFS [78] as file system for the images, which is read-only and compressible. For distribution of the images, CERN hosts a CVMFS repository `/cvmfs/unpacked.cern.ch`. It provides the DOCKER and SINGULARITY images in uncompressed form, which can then be accessed anywhere. Using CVMFS for the distribution, users can benefit from the advanced caching setup, described above.

Another type of image are called *sandbox* images. These are basically the files of generated image within a directory structure. This type of image has the advantage of remaining writable if the underlying file system is writable, allowing to change images from within a running container. Such sandbox images are mostly used for development, however, as the underlying file system is writable they should normally not used for production.

6.4.3. User Namespaces

SINGULARITY requires specific privileges on the host system to start up a container from a given image file, however the privileges within the container are limited to what the linux user is authorized to do on the host system. Because of this, running a SINGULARITY container within another SINGULARITY container is not possible trivially but requires some extra steps.

One possibility is to start up the outer container as the user `root`, which has full privileges on the host system. This, however, besides not being possible on systems like HPC clusters and sites operated by third-parties, is not desirable as it would allow for privilege escalation from within the container. To remedy such problems, the linux kernel lets users and processes create namespaces. A namespace in this context enables processes to see only a limited set of system resources like files, other processes or user IDs. With this, processes within a namespace can be denied specific system resources while still being able to access others. More specifically, this makes it possible to enable a process the access to system resources it normally wouldn't have, without granting full privileges. [79]

Namespaces are grouped in different kinds, depending on the resources they isolate for the process. *User namespaces* map user and group IDs from within the namespace to the host system. This allows to run processes within the namespace with a different user ID than the one on the host system. It is for example possible to have a user with ID `0` within the namespace, which can act as user `root`, but only in the constrained realm of the namespace. This is useful as the user has virtually administrative privileges within the namespace.

Jobs within the WLCG as introduced in section 2.1 are started within a container by the batch system they are managed with. As the external sites which are integrated using COBALD/TARDIS should also be able to run pilot jobs from the WLCG, but the OBS which manages the integrated resources starts its payloads in turn itself within a container, the use of user namespaces enables this mode of operation.

6.5. Grid Site Operation and Remote Compute Element

WLCG grid sites serve experiments for their event reconstruction, MC simulations, and analyses. To be able to operate the site and run such workflows, besides CPU and storage resources, several other services have to be provided. These include but are not limited to a batch system for managing the compute resources (the LRMS), authorization and accounting of users as well as experiments, monitoring, CVMFS as discussed above, possibly a local squid proxy for caching, and a compute element (CE).

Instead of regarding a grid site a monolithic entity it can be considered a set of modules, which in parts can be operated separately from each other. In order to enable smaller sites to contribute computing power to the WLCG, some of the components might be able to be offered as a service by other regional sites. This takes some workload of the local team and allows them to devote their time towards the remaining tasks.

The compute element is the gateway to grid sites for users and experiments. It acts as entry point for them and allows to submit jobs to the affiliated site from remote. For experiments, the use of a gateway software for accessing the resources, eases their site management, as only a handful of different types of gateways exists. The CE checks the authorization of the users and their experiment affiliation and whether it is authorized to use the resources, allowing it to forward the jobs to specific resources, the users of the experiment have access to. For the authentication X.509-certificates, as discussed in the next section, are used. [80]

The CE forwards the jobs to the LRMS by submitting copies of the jobs to the corresponding queues of the LRMS of the site, from where they are matched to the specific resources. As the CE and the LRMS are disjoint, the former can be operated independently of the latter. This also allows to separate the CE from the site it services and to implement it at a different location. With this it was possible to deploy a remote CE for the Aachen tier 2 WLCG site. While all other services, especially computing and storage resources remain in Aachen, the operation of the CE of the site was moved to the Karlsruhe tier 1 WLCG site. The Karlsruhe GridKa team already operates several of such CEs for the local tier 1. At the tier 1, the software used for the CE is the HTCONDOR-CE. Deployment and configuration of them is completely automated, using the software configuration management tool PUPPET [81]. The overhead of adding and managing an additional CE is hence negligible. This project allows the Aachen grid team to discontinue the operation of the local CE, focusing on the operation of the other core services.

6.6. Access Rights and Site Policies

Every site has to authenticate their users, to check if they are who they claim to be and can be handled differently, depending on the intended mode of access to the resources. Most sites aim at users using their account, registered locally at the specific site, exclusively for their own workflows. This is especially true in the HPC context. With this model, authentication can be done by the user providing a username and password, or some other proof of identity like an SSH-key or access token. This model, however, fails, if such mapping is not possible as the single users do not register with the site itself. For this, authentication models exist which allow for the identification to be performed by a third party.

With a mode of operation as implied by the use of COBALD/TARDIS, this kind of authentication of single users directly on site is not possible. Instead, the resource allocation is done using a user account, registered at the site. This user account serves as a proxy to be able to access the resources of the site and in terms of accounting represents the research groups associated to the project.

Authentication in this case is performed against the OBS. This authentication and use of the resources of the site though implies a certain level of trust between the operators of the site and the OBS. The usage policy of the site, the set of rules which defines how the site is allowed and intended to be used, defines whether this type of trust of third parties can be established or not.

Many sites define policies which explicitly disallow to relay user accounts to others. As this specifically prohibits the use of the resources in a mode as intended with COBALD/TARDIS, exemptions have to be negotiated with the operators of the site for their integration. Strong and tight communication with the site operators is crucial for such negotiations. This is a lot easier for groups, local to the site, i.e. research groups from the university or research institution. Thus, it is beneficial to have local groups work on the integration of their local sites into a distributed system like the one built with COBALD/TARDIS.

In the grid-context, especially within the WLCG, authentication and authorization is done very differently. Each user has a personal X.509-certificate [82] for authentication. These certificates are issued by centrally approved certification authoritys (CAs) which act as trust entities. A CA is allowed to issue certificates to users, after the user has proven their identity to the CA. With this approach a level of trust can be built which enables the usage of sites for users, which they are completely alien to.

Furthermore, the user can register their certificate with the virtual organization of their experiment through the virtual organization membership service (VOMS). This allows sites to identify the experiment affiliation of the user and thus enables more finely grained access controls. The issued certificates normally have lifetimes of a year.

For normal everyday use, ephemeral proxies with lifetimes in the order of hours or days are derived from the certificate. Generating the proxy, the affiliation to the experiment or even specific roles within the experiment are checked through the VOMS and added to the proxy as an extension. The user afterwards submits the proxy together with their jobs, which in turn use it to identify the user on the grid. Compute and storage providers use the provided proxy to authenticate the users and check their authorizations. As the proxy can act as an unambiguous identification of the user and its affiliation to experiments, policies can be implemented at each level, granting or denying the user and their jobs access to specific resources. Through the use of the extensions, even distinguishing roles can be given, granting access to specific resources.

The policies of how a resources may be used and which user is allowed to use it differs widely as mentioned above. Within the WLCG, resource usage is enabled through the just mentioned proxies with their possible extensions, which are checked are the WLCG sites. Computing resources in the university environment on the other hand are far more heterogeneous and often closely coupled to local users of the university or local computing projects. Very tight restrictions on the use of the resource is often owed to the funding agencies.

This is a challenge, especially in Germany: Because of the federated structure of Germany, funding situations are often complicated. Education, which include also university matters, are governed on state level (*Länder*). Collaborations like the one from CMS, however, are managed on federal level (*Bund*). Because science, including HEP experiments as well as the official WLCG computing resources are funded on federal level and pledged to the experiments, those can be and are available to the collaborations. Computing resources funded by the states on the other hand, are operated for the local universities and the local research groups. Because of this, the use of them is more restricted and there is no explicit affiliation of the computing resources to experiments.

To make state resources, however, available for users through decentralized infrastructures, which serves users not only from this specific state, either restrictions have to be implemented on OBS level, to only authorize the use of the resources for the allowed users, or exemptions have to be negotiated with site operators and possibly state governments. These exemptions should be regarded with the benefits, the local users can have from collaborating with other research groups from within the experiment, including on workflow steering level, is paramount to maintaining competitiveness.

6.7. Backfilling

Usual jobs on HPC systems often require a large number of resources, sometimes not available with just one worker node, thus spanning several nodes. As the next job in the queue can only be matched to its resources once they are free, the resources required for it are reserved and as soon as all previous jobs on these resources have finished, this job can be matched. This is done by not starting new jobs on the reserved resources, even though some job might fit. The reserved resources are "waiting" until enough resources are available, so to fit the larger job.

While the resources are not matched with other jobs, waiting for the larger job to start, they idle. As resources are not matched with jobs for possibly quite some time, these resources are not available for other jobs, this can waste a lot of compute time. This in turn can degrade the utilization of the site, often a key indicator, when arguing with funding agencies.

Specific types of jobs can take advantage of this and fill up the temporarily idling resources. By matching jobs from a certain queue to the resources waiting to be matched with larger jobs, the time resources go to waste can be minimized. One prerequisite for these jobs, however, is that they have to finish before the larger jobs wants to claim the resources. If running longer than the allotted time the job has to be able to be removed from the resources at any moment. Such a mode of operation of a batch system is called *backfilling*.

If a backfilled job is removed from the resource, by just killing the process it normally looses all its progress. The job needs to be put back in the queue and will be matched again when appropriate resources become available. Depending on the batch system and the job, before hard killing the latter, it can be given a signal by the batch system. The job then has the possibility to save intermediate results, or write out a checkpoint file. After some grace period, the job is killed. Jobs, able to save checkpoint files, can be restarted again, resuming computation from their last checkpoint.

With this, the utilization of the resource can be increased greatly, while providing additional, opportunistic compute resources for other scientific communities. In HEP, workflows can be defined very granular: Often the same executable is used to process hundreds, or even thousands of the same type of input file. Their runtime ranges in the order of up to a few hours, which suits very well for the backfilling.

6.8. Opportunistic Extension of the Aachen Tier 2 WLCG Site

The research groups at the RWTH Aachen University have access to the high-performance computing (HPC) cluster CLAIX [83]. The system provides roughly 77 000 CPU cores and several parallel cluster file systems, all connected by the high performance network architecture Omni-Path [84]. Access to the cluster is granted to members and partners of RWTH on a small scale. Researchers from other research institutes or groups with a high demand in computing time can, however, submit applications for larger computing projects.

The cluster uses CentOS in version 7.9 as operating system. It can be accessed via SSH [85] through a set of login-nodes. These are used for the submission of the jobs to the LRMS SLURM, which takes care of the resource management.

With this, CLAIX is perfectly suited to be used for HEP workflows:

1. It is located at a university at which the physics department operates a tier 2 WLCG site, hence it has local experts for close communication and with a strong background in grid computing.
2. The network connection between CLAIX and the tier 2 is extremely good, the data storage infrastructure for many HEP experiments is available as performant as if it was local to CLAIX.
3. The cluster already acts as a CVMFS client and enables the usage of SINGULARITY containers with user namespaces.
4. The resources can be accessed via SSH and are plentiful.

This list fulfills all requirements stated in section 6.3.

With this positive assessment it was decided to dynamically extend the tier 2 WLCG site with the opportunistic integration of CLAIX. For this, a COBALD/TARDIS instance was deployed handling the scheduling of the resources and their life cycle management. The HTCONDOR instance, operated at the physics department of the RWTH, which manages the

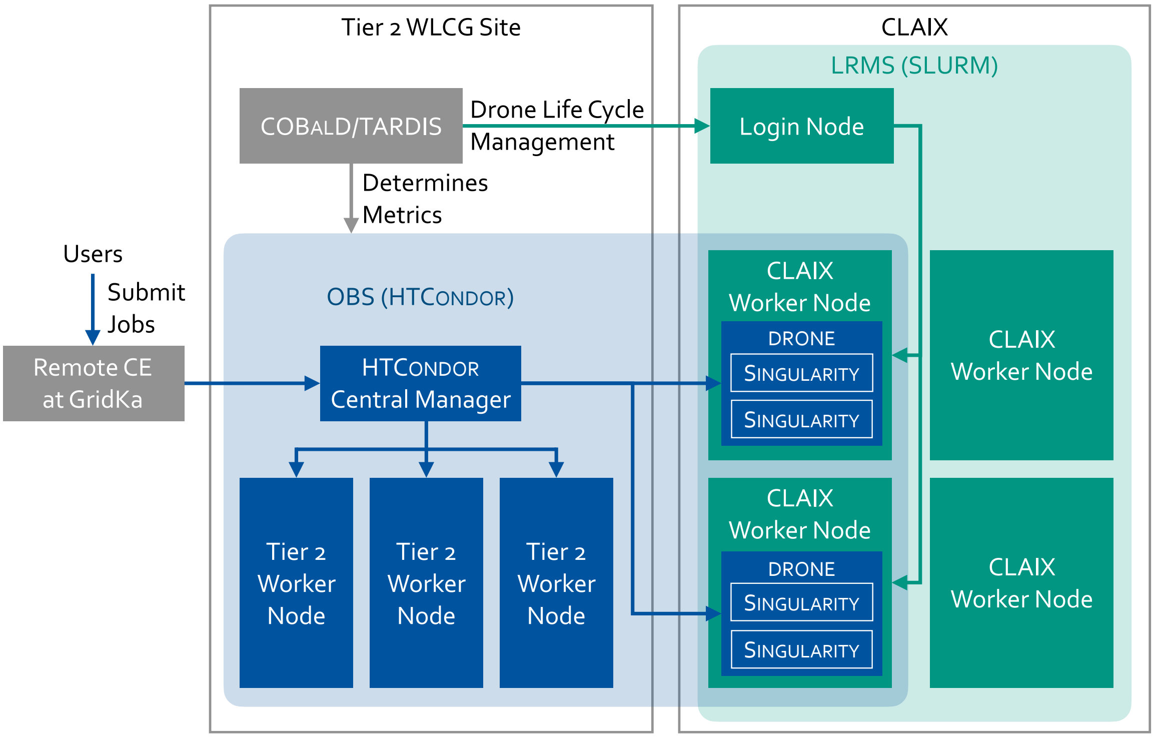

Figure 6.4.: The COBALD/TARDIS setup used for integrating resources of the HPC cluster CLAIX into the RWTH tier 2 WLCG site to make them available for users and experiments. The flow of the jobs is indicated by blue arrows, the life cycle management of the DRONES by COBALD/TARDIS is indicated by green arrows. The blue boxes within the shaded blue area are managed by the OBS HTCONDOR, while the green boxes within the shaded green area are managed by CLAIX' LRMS SLURM.
Users and experiments submit their (pilot) jobs to the HTCONDOR-CE, which dispatches them to the HTCONDOR-OBS. The OBS decides onto which resource to match the job. COBALD/TARDIS periodically checks for the metrics occupancy and suitability and increases or decreases the number of DRONES at CLAIX, as necessary by submitting or cancelling DRONES through CLAIX' LRMS. The LRMS matches the pilot-jobs to the worker nodes of CLAIX. After the DRONES start up, they connect to the central manager of the OBS and are ready to accept jobs. Any job matched to a DRONE is then run within a SINGULARITY container.

resources of the tier 2 is used as the OBS. With this integration and operation mode, the HPC cluster CLAIX is available to run official CMS workflows and user jobs.

The allocation of the resources is performed by submitting DRONES to the LRMS of the site. Authentication of the local employee user account, used for the submission of the DRONES is done using an encrypted SSH-key on the service machine, where COBALD/TARDIS is installed. The DRONE consists of a bash script, the HTCONDOR software as precompiled binaries and the corresponding configuration files. Furthermore it has access to the password of the HTCONDOR instance (pool password) acting as OBS to authenticate any HTCONDOR-daemon.

When the DRONE boots up on the worker node of the cluster, it sets up a temporary directory for the OBS to run in. It duplicates the HTCONDOR binaries onto the node and copies over the provided configuration files. Further, dynamically generated configuration files are added, necessary for the OBS to function. The DRONE then starts HTCONDOR, which

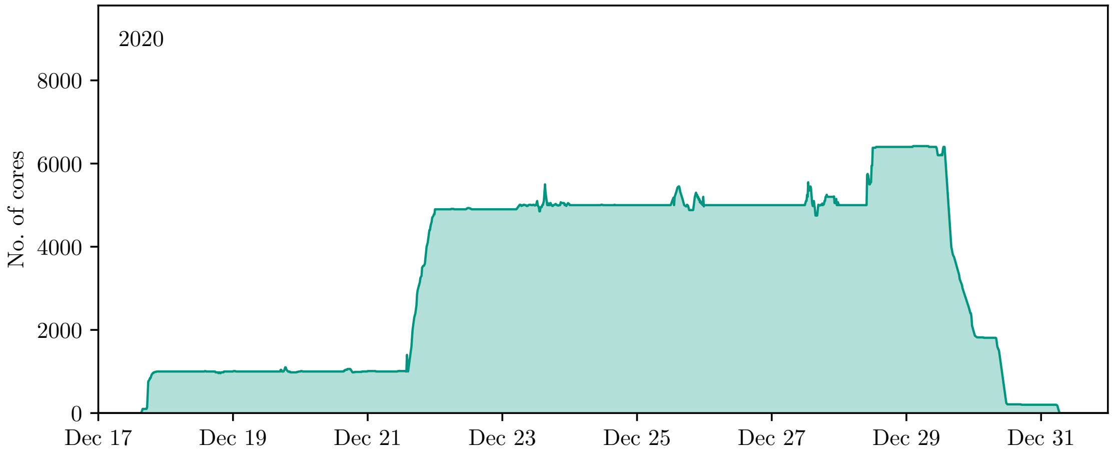

Figure 6.5.: The number of cores running on CLAIX integrated into the OBS operated at the tier 2 WLCG site at RWTH, during the Christmas break of 2020. This was accomplished within the pilot project.

is configured to run the `condor_startd` on the node. At this point, the DRONE is ready to accept jobs from the OBS. An overview of the setup is shown in fig. 6.4.

The jobs running on the CLAIX resources are started in SINGULARITY containers, specifying the environment expected at WLCG resources. The "Dockerfile", a configuration file describing the image of the container is given in reference [86]. The image is provided using the `/cvmfs/unpacked.cern.ch` repository already configured on the nodes. When starting the container the CVMFS repositories from the worker, `/etc/hosts` and `/etc/localtime` are bindmounted into the container. The latter two to provide name resolution and timezone information.

As the resources are integrated to be part of a WLCG site, for CMS jobs some additional information has to be provided. This is done through the directory `/cvmfs/cms.cern.ch/SITECONF/local`. On the worker node, this is configured to be a symlink to the path `/etc/cvmfs/SITECONF`, however, this directory does not exists on the host. When starting the container, the respective configuration directory is bindmounted to exactly this path, providing the necessary information at the right place accessible for the jobs within the container.

The jobs being processed on the tier 2 WLCG site are often official CMS workflows. CMS does its resource allocation in a similar manner, as the COBALD/TARDIS setup. By submitting pilot-jobs to the individual sites (in this case the numerous WLCG sites), the resources are integrated into yet another batch system, operated by the experiment. The payloads of this experiment batch system are again started in SINGULARITY containers in order to provide the expected experiment software and environment. Because of this it has to be possible to start SINGULARITY containers within a SINGULARITY container. As discussed above in section 6.4.3 this is possible with the use of user namespaces. So for the setup used at CLAIX, the containers started by the OBS operated by the tier 2 WLCG site are configured to run with user namespaces.

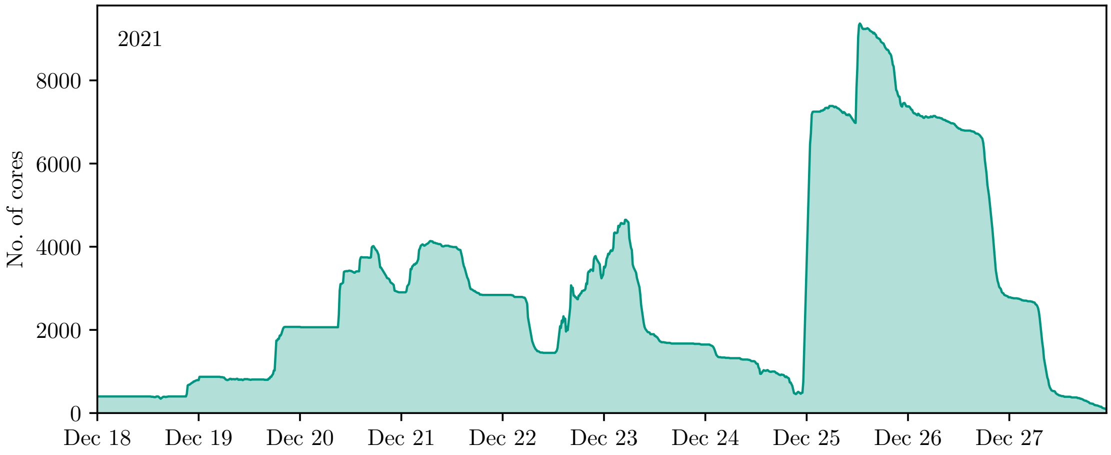

Figure 6.6.: The number of cores, integrated from the HPC cluster CLAIX into the tier 2 WLCG site at RWTH, from the production phase in year 2021.

With this setup, it was possible to transparently integrate 5000 cores for use in the tier 2 WLCG site and provide these resources to the CMS experiment. This was done as a pilot project to prototype and proof the feasibility of the setup. In fig. 6.5 the number of cores, integrated into the tier 2 WLCG site is shown, over the span of the pilot project in December 2020. After a setup and configuration phase of a couple of days, the system was scaled up to integrate 1000 cores for testing, before ramping up to the peak 5000 cores. Without any modifications, the same job mix, also processed on the official tier 2 resources were run on the CLAIX resources. This was only possible because of the excellent network connection of CLAIX to the tier 2 WLCG site, where experiment data storage is located. After roughly eight days of usage, further submission got blocked by the LRMS of the HPC cluster, as the eightfold of the allotted computing time was used within this time span.

The periodic modulations, noticeable especially on December 23, 25 and 27, are because of the two-days lifetime of the DRONEs and all of them having been started nearly at the same time. The period with more than 5000 integrated cores, was an issue in translating states of the jobs within the LRMS to the states of the DRONEs within COBALD/TARDIS. This was revised afterwards. Within the shown 15 days, CLAIX resources have provided 1.13 million core-hours for the CMS experiment through the WLCG.

Such a use of the HPC cluster CLAIX is, however, a breach of the *RWTH Netzordnung*, the usage policies of their computing infrastructure. As the jobs get matched onto the resources by the OBS, they can be submitted by any user of the involved VOs. However, as the research groups in this project directly profit from the jobs processed by the tier 2 infrastructure, and the submitter from each job can be traced back because they authenticated with their X.509-certificate against the OBS, an exemption to the policy could be negotiated with the operators of CLAIX.

Having proven the possibility of using the HPC cluster CLAIX resources as part of an official tier 2 WLCG site, a computing project to move the setup to production was applied

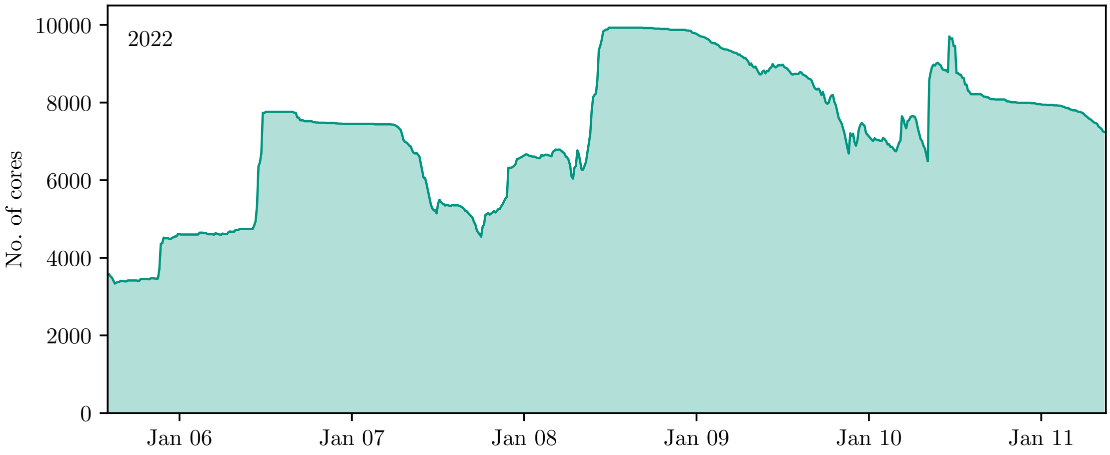

Figure 6.7.: The number of cores, integrated from the HPC cluster CLAIX into the tier 2 WLCG site at RWTH, in January 2022.

for. The new computing project was granted 7.5 million core-hours for one year. This allows to dynamically extend the WLCG site with roughly 850 cores on average. In fig. 6.6 the number of cores, integrated into the tier 2 WLCG site, is shown over nine days of December 2021. Within the week of December 20 to December 27, 2021, 0.64 million core-hours were provided.

The HPC cluster CLAIX continues to deliver outstanding performance: within just less than six days in January 2022, it provided 1 million core-hours for the CMS experiment to use. This is shown in fig. 6.7, note the change in the y-axis compared with figs. 6.5 and 6.6. This is equivalent to 7300 cores on average, which, temporarily, more than triples the amount of resources provided to the WLCG through the Aachen tier 2 WLCG site.

CLAIX is the first HPC cluster, which is integrated into a German tier 2 WLCG site. Its successful integration proves the possibility of using HPC clusters as extensions for WLCG centers. The outstanding performance demonstrates the potential, non-HEP sites have in providing computing resources for HEP experiments. Reaching agreements with site operators to use the resources with a computing model as the one presented here, will allow to continue delivering an substantial amount of resources to the WLCG despite the expected change in the funding situation in Germany in the future.

7. Conclusion and Outlook

In the scope of this thesis, the derivation of the absolute residual corrections using Z + jet events as part of the intricate process of correcting the jet energy scale at the CMS detector has been discussed and performed with the data recorded in the year 2018. Furthermore, a modern, dynamic way of integrating additional non-HEP computing resources into existing computing infrastructures was described, allowing a transparent usage of these resources for both, users and experiments.

The jet energy corrections are derived with the data recorded in the data-taking period of 2018 with an integrated luminosity of $\mathcal{L} = 59.83\,\mathrm{fb}^{-1}$ of proton-proton collisions at a center-of-mass energy of 13 TeV as laid out in chapter 5. The correction factors were extracted as part of the elaborate workflow, employed at CMS for the jet energy calibration. For the derivation of the absolute residual corrections, events containing a jet in addition to a pair of electrons or muons, from which a Z boson can be reconstructed, are selected. The kinematic properties of the event allow to balance the transverse momentum of the jet against the transverse momentum of the reconstructed Z boson.

The level 3 absolute residual corrections are derived to alleviate differences in the measured jet energy scale after mitigating contributions to the measurements from pileup (PU) and discrepancies between the reconstruction and generator level of simulated samples. Two different methods based on two observables, p_T-balance and missing transverse energy projection fraction (MPF), are used for this. While both methods balance the jet against the Z boson as reference object and use the kinematic properties of the two objects to be of opposite transverse momentum, the MPF method additionally accounts for the missing transverse momentum in the events. Both methods are based on an event topology without additional jets. The responses derived with the methods are thus prone to further jets in the event, quantified by the *additional jet activity* α. By deriving the responses in bins of α, the correction factors can be extrapolated to show no additional jet activity $\alpha \rightarrow 0$ in the event.

The responses derived with the analysis of the Z + jet events are then combined with similar analyses using γ + jet, and multi-jet events as well as the two jets from a hadronically decaying W boson from semileptonic $t\bar{t}$ events in a global fitting procedure performed by the *jet energy resolution and correction* group of the CMS experiment. The resulting correction factors will be used in future analyses, performed on the data recorded in the year 2018.

Calibration efforts for the data recorded in years 2016 and 2017 with the CMS detector and the corresponding publication on the jet energy calibrations, including all data taken during Run II of the LHC, are still ongoing.

For carrying out such calibration tasks, and to perform physics analyses like cross-section measurements, or searches for new physics, an enormous amount of computing resources is needed. For this, the WLCG provides scientists and experiment collaborations with the required computing and storage facilities. With this global network of computing centers, reconstruction, analysis, and simulation workflows can be performed. Studies predict, however, a shortage of CPU resources of up to a factor of four by the year 2034 in the worst scenario, as stated in chapter 2.

Besides further improving the algorithms and software, used for processing the data, and advancing and deploying novel hardware, one way of alleviating this shortage is the utilization of additional resources from outside the community. Such resources may include, for example, high-performance computing (HPC) or university clusters, as well as science or commercial clouds. A way for users and experiments to transparently access these resources has to be established. Furthermore, the allocation of the resources has to be automated, making them available whenever necessary.

All this is accomplished with the use of the resource scheduler COBALD/TARDIS, using a generalization of the pilot concept, as established within the WLCG. COBALD/TARDIS integrates resources on demand into one common overlay batch system (OBS), acting as a single point of entry. Instead of accessing each resource individually and submitting the computing payloads as jobs to the individual local resource management systems (LRMSs), users and experiments submit their jobs to the OBS, authenticating only once. The OBS takes care of dispatching the jobs to the integrated resources. By utilizing the OBS to determine metrics of occupancy and suitability of the sites, COBALD/TARDIS can make scheduling decisions using a feedback loop, reacting to the decisions of the OBS, this allows for dynamic resource allocation based on when they are needed for the current job mix. This increases the number of well-utilized resources, improving the overall utilization of the site.

Non-community specific resources, however, often lack particular software or protocols for access to certain storage or file systems. This has to be provided so that the resource can be used transparently. It can be achieved by executing the computing payloads of the jobs which are submitted to the OBS, either in a containerized environment or a virtual machine.

With a COBALD/TARDIS setup and the use of SINGULARITY containers it was possible to dynamically and transparently extend the Aachen WLCG tier 2 site, using resources provided by a HPC cluster operated by the local university. The HTCONDOR batch system serving the tier 2 site, operated by the physics department of the RWTH Aachen University, was used as the OBS to integrate the resources from HPC cluster CLAIX, which in turn is operated by the IT department of the university. By running a COBALD/TARDIS instance, the resources were dynamically integrated into the OBS, allowing any job scheduled to run on the tier 2 site, to be matched against resources on the HPC cluster and executed there. The resource allocation is performed by submitting placeholder jobs to the LRMS of the HPC cluster. The jobs matched to the HPC resources are configured to run in SINGULARITY containers, which provide the expected WLCG environment. The setup was configured to enable the use of nested SINGULARITY containers, which is required to run official CMS campaigns on the resource. Within a pilot project, it was possible to allocate around 5000 CPU cores for more than a week. After the successful test phase the project was moved into production mode, providing additional 7.5 million core-hours over one year.

The cluster continued to deliver excellent performance, contributing substantial computing resources. Within less than a week, one million core-hours were provided to the CMS experiment through the Aachen WLCG tier 2 site. This temporarily tripled the amount of computing power, provided by this single site. Its successful integration proves the practicability of using non-HEP resources to mitigate resource shortages, expected in the future.

For a streamlined entry point to grid resources, each grid site has to operate a compute element (CE), which takes care of authentication, accounting, and forwarding of the jobs to the underlying LRMS of the resources. Within the projects of this thesis, a “remote” CE was set up at GridKa in Karlsruhe, serving the WLCG tier 2 site in Aachen. Successfully compartmentalizing grid sites and therefore being able to have single parts operated as remote services by third parties, facilitates the operation of a such a grid site enormously.

A setup with remote grid services, or integrated with another grid site using OBSs, enables smaller sites to contribute computing power, which benefits the whole community. In such a distributed computing infrastructure, which allows users to utilize virtually any resource, tools will have to be developed to take care of the accounting and reporting of which user groups use which resources. This is especially important in e.g. the German computing landscape, in which the funding situation of many projects is often highly complex. Furthermore, with the improvements to network infrastructure which are anticipated for the near future, new computing models will emerge and benefit from the abundance of resources becoming available. It is expected that one improvement to the network will be the connection of different sites, which will make the operating mode as presented here even more reliable for data-intensive workflows such as particle physics analyses.

Both parts of this thesis, the jet energy calibration and the dynamic integration of heterogeneous computing resources, provide an important contribution to published and future physics analyses. First, the measurements of the absolute residual corrections using Z + jet events allow to precisely calibrate jets in the data taken in year 2018 with the CMS detector. With calibrated jets, precision measurements of cross-sections or searches for new physics are made possible. Second, the developments done in the scope of dynamic resource allocation in a heterogeneous computing landscape allow for the transparent usage of non-community resources for HEP workflows, alleviating expected computing resource shortages in the future.

A. Appendix

Table A.1.: The data samples used in this calibration, recorded with the CMS detector at $\sqrt{s} = 13\,\text{TeV}$ in 2018. The integrated luminosity of the data samples is $59.83\,\text{fb}^{-1}$. The type indicates the selection of trigger path of the samples.

Type	Version	Era	No. of events
EGamma	12Nov2019_UL2018-v2	A	339 013 231
		B	153 822 427
		C	147 827 904
	12Nov2019_UL2018-v4	D	752 489 952
DoubleMuon	12Nov2019_UL2018-v2	A	75 491 789
		B	35 057 758
		C	34 565 869
	12Nov2019_UL2018-v3	D	168 552 341

Table A.2.: The simulated data sample used in this calibration. The events are simulated using PYTHIA8 [30] and MADGRAPH [42].

Name	No. of events
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunII Summer19UL18MiniAOD-106X_upgrade2018_realistic_v11_L1v1-v1	104 017 741

Table A.3.: The CMS PU jet identification for AK4CHS jets for the data recorded in the year 2018, reconstructed with the ultra legacy campaign.

(a) Tight working point

	$\lvert\eta\rvert < 2.5$	$2.5 < \lvert\eta\rvert < 2.75$	$2.75 < \lvert\eta\rvert < 3.0$	$3.0 < \lvert\eta\rvert < 5.0$
$10 < p_T < 20$	0.77	0.38	−0.31	−0.21
$20 < p_T < 30$	0.90	0.60	−0.12	−0.13
$30 < p_T < 40$	0.96	0.82	0.20	0.09
$40 < p_T < 50$	0.98	0.92	0.47	0.29

(b) Medium working point

	$\lvert\eta\rvert < 2.5$	$2.5 < \lvert\eta\rvert < 2.75$	$2.75 < \lvert\eta\rvert < 3.0$	$3.0 < \lvert\eta\rvert < 5.0$
$10 < p_T < 20$	0.26	−0.33	−0.54	−0.37
$20 < p_T < 30$	0.68	−0.04	−0.43	−0.30
$30 < p_T < 40$	0.90	0.36	−0.16	−0.09
$40 < p_T < 50$	0.96	0.61	0.14	0.12

(c) Loose working point

	$\lvert\eta\rvert < 2.5$	$2.5 < \lvert\eta\rvert < 2.75$	$2.75 < \lvert\eta\rvert < 3.0$	$3.0 < \lvert\eta\rvert < 5.0$
$10 < p_T < 20$	−0.95	−0.72	−0.68	−0.47
$20 < p_T < 30$	−0.88	−0.55	−0.60	−0.43
$30 < p_T < 40$	−0.63	−0.18	−0.43	−0.24
$40 < p_T < 50$	−0.19	0.22	−0.13	−0.03

Acronyms

ALICE A Large Ion Collider Experiment

AOD Analysis Object Data

ATLAS A Toroidal LHC ApparatuS

BOOSTER Proton Synchrotron Booster

CA certification authority

CE compute element

CERN European Organization for Nuclear Research

CERN Organisation européenne pour la recherche nucléaire

CHS charged hadron subtraction

CMS Compact Muon Solenoid

CSC cathode strip chamber

CVMFS CERN virtual machine file system

DAQ data acquisition

DESY Deutsches Elektronen-Synchrotron

DT drift tube

EB ECAL barrel

ECAL electromagnetic calorimeter

EDM Event Data Model

EE ECAL endcap

ES ECAL preshower

ETP Institute of Experimental Particle Physics

EW electroweak theory

FSR final state radiation

GCT global calorimeter trigger

GMT global muon trigger

GPU graphics processing unit

GT L1 global trigger system

HB HCAL barrel

HCAL hadronic calorimeter

HE HCAL endcap

HEP high energy physics

HEP-SPEC06 HEP Standard Performance Evaluation Corporation 06

HF HCAL forward

HLT high level trigger

HO HCAL outer

HPC high-performance computing

HTTP hypertext transfer protocol

IP interaction point

ISR initial state radiation

JDL job description language

JEC jet energy calibration

JERC *jet energy resolution and correction*

KIT	Karlsruhe Institute of Technology
L1	level 1 trigger
L1A	L1 accept
LEP	Large Electron-Positron Collider
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LHCf	Large Hadron Collider forward
LHCONE	LHC Open Network Environment
LHCOPN	LHC Optical Private Network
LRMS	local resource management system
LV	leading vertex
MC	Monte Carlo
MET	missing transverse momentum
MFA	multi-factor authentication
MINIAOD	Mini AOD
MPF	missing transverse energy projection fraction
NAF	National Analysis Facility
NANOAOD	Nano AOD
OBS	overlay batch system
OS	operating system
OTP	one-time password
PDF	parton distribution function
PF	particle flow
PS	Proton Synchrotron
PU	pileup
PUPPI	pileup per particle identification
QCD	quantum chromodynamics
RCT	regional calorimeter trigger
RHEL	Red Hat Enterprise Linux
RPC	resistive plate chamber
RWTH	RWTH Aachen University
SM	Standard Model
SPS	Super Proton Synchrotron
TCS	trigger control system
TEC	tracker end caps
TIB	tracker inner barrel
TID	tracker inner disk
TOB	tracker outer barrel
TOTEM	Total Elastic and Diffractive Cross Section Measurement
UHH	Universität Hamburg
VM	virtual machine
VO	virtual organization
VOMS	virtual organization membership service
WLCG	Worldwide LHC Computing Grid

List of Figures

2.1. The structure of the WLCG 6
2.2. The CPU performance in kHEP-SPEC06 and disk storage in petabyte the CMS collaboration expects to require per year for up to 2034 as of November 2021 7

3.1. The fundamental vertices of quantum chromodynamics 11
3.2. Feynman diagrams of the leading order production processes of a Z boson in association with a quark or gluon and the Z boson decaying into a pair of electrons or muons 13

4.1. The LHC accelerator complex 16
4.2. CMS experiment with its subdetectors 17
4.3. The coordinate system used at CMS 18

5.1. Examples for a collinear and infrared unsafe jet clustering 27
5.2. The factorized workflow of the jet energy calibration 30
5.3. Distributions of the transverse momentum of the leading jet 37
5.4. Distributions of the pseudorapidity of the leading jet 38
5.5. Distributions of the transverse momentum of the reconstructed Z boson . . . 38
5.6. Distributions of the mass of the reconstructed Z boson 39
5.7. Distributions of the missing transverse momentum in the events 40
5.8. Responses derived with the MPF method on data taken in year 2018 40
5.9. Responses derived with the p_{T}-balance method on data taken in year 2018 . . 41
5.10. Ratio of the jet response between data and MC simulations after the global fitting procedure for year 2018. 42
5.11. Distributions of the MPF responses as a function of the transverse momentum of the Z boson 43
5.12. Distributions of the MPF responses as a function of the pseudorapidity of the leading jet 44
5.13. Ratios of the mean values of the MPF response in data over MC as a function of the transverse momentum of the Z boson and the leading jet pseudorapidity 45
5.14. Ratios of the mean values of the p_{T}-balance response in data over MC as a function of the transverse momentum of the Z boson and the leading jet pseudorapidity 46

6.1. Depiction of the metrics occupancy and suitability 51
6.2. Feedback loop, COBALD/TARDIS uses to schedule resources 52
6.3. Virtualization technologies 56

6.4. The COBALD/TARDIS setup used for integrating resources of the HPC cluster CLAIX into the RWTH tier 2 WLCG site. 62
6.5. The number of cores running on CLAIX integrated into the OBS operated at the tier 2 WLCG site at RWTH, during the Christmas break of 2020 63
6.6. The number of cores, integrated from the HPC cluster CLAIX into the tier 2 WLCG site at RWTH, from the production phase in year 2021 64
6.7. The number of cores, integrated from the HPC cluster CLAIX into the tier 2 WLCG site at RWTH, in January 2022 65

References

[1] David Griffiths. *Introduction to elementary particles*. 2008. ISBN: 978-3-527-40601-2.

[2] *Convention for the establishment of a European organization for nuclear research: Paris, 1st July, 1953 : as amended. Convention pour l'établissement d'une Organisation européenne pour la Recherche nucléaire. Paris, le 1er juillet 1953 : telle qu'elle a été modifiée*. Geneva: CERN, 1971. URL: https://cds.cern.ch/record/330625.

[3] *The Synchrocyclotron*. CERN. URL: http://home.cern/science/accelerators/synchrocyclotron (visited on 01/02/2022).

[4] K. Bos et al. *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical design report. LCG. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/840543.

[5] I Bird et al. *Update of the Computing Models of the WLCG and the LHC Experiments*. Tech. rep. Apr. 2014. URL: https://cds.cern.ch/record/1695401.

[6] *WLCG – Thumbnails*. July 2014. URL: https://espace.cern.ch/WLCG-document-repository/images1/Forms/Thumbnails.aspx?RootFolder=%2FWLCG-document-repository%2Fimages1%2FWLCG (visited on 09/06/2021).

[7] HEPiX working groups. *Benchmarking Working Group*. URL: http://w3.hepix.org/benchmarking (visited on 09/08/2021).

[8] CMS Collaboration. *CMS Offline and Computing Public Results*. Nov. 2021. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults?rev=14 (visited on 12/23/2021).

[9] M Ernst et al. *dCache, a distributed storage data caching system*. July 2001.

[10] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[11] M. C. Gonzalez-Garcia and Yosef Nir. "Neutrino masses and mixing: evidence and implications". In: *Reviews of Modern Physics* 75.2 (Mar. 2003), pp. 345–402. ISSN: 1539-0756. DOI: 10.1103/revmodphys.75.345. URL: http://dx.doi.org/10.1103/RevModPhys.75.345.

[12] A. D. Martin et al. "Parton distributions for the LHC". In: *The European Physical Journal C* 63.2 (July 2009), pp. 189–285. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-009-1072-5. URL: http://dx.doi.org/10.1140/epjc/s10052-009-1072-5.

[13] Klaus Rabbertz. *Jet Physics at the LHC*. Vol. 268. Springer Tracts in Modern Physics. Cham: Springer International Publishing, 2017. ISBN: 978-3-319-42115-5. DOI: 10.1007/978-3-319-42115-5. URL: http://link.springer.com/10.1007/978-3-319-42115-5 (visited on 11/14/2021).

[14] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[15] Murray Gell-Mann. "Symmetries of Baryons and Mesons". In: *Phys. Rev.* 125 (3 Feb. 1962), pp. 1067–1084. DOI: 10.1103/PhysRev.125.1067. URL: https://link.aps.org/doi/10.1103/PhysRev.125.1067.

[16] G. Arnison et al. "Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider". In: *Physics Letters B* 126.5 (1983), pp. 398–410. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90188-0. URL: https://www.sciencedirect.com/science/article/pii/0370269383901880.

[17] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088/1748-0221/3/08/s08001.

[18] Esma Mobs. "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018". In: (Aug. 2018). General Photo. URL: https://cds.cern.ch/record/2636343.

[19] The LHCf Collaboration. "The LHCf detector at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08006–S08006. DOI: 10.1088/1748-0221/3/08/s08006. URL: https://doi.org/10.1088/1748-0221/3/08/s08006.

[20] The TOTEM Collaboration. "The TOTEM Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08007–S08007. DOI: 10.1088/1748-0221/3/08/s08007. URL: https://doi.org/10.1088/1748-0221/3/08/s08007.

[21] CMS Collaboration. "The CMS experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[22] *latex:example_spherical_coordinates [CMS Wiki Pages]*. URL: https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates (visited on 12/22/2021).

[23] Tracker Group of the CMS Collaboration. "The CMS Phase-1 pixel detector upgrade". In: *Journal of Instrumentation* 16.02 (Feb. 2021), P02027–P02027. DOI: 10.1088/1748-0221/16/02/p02027. URL: https://doi.org/10.1088/1748-0221/16/02/p02027.

[24] M. Abbrescia et al. “Beam test results on double-gap resistive plate chambers proposed for CMS experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 414.2 (1998), pp. 135–148. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(98)00571-3. URL: https://www.sciencedirect.com/science/article/pii/S0168900298005713.

[25] CMS Collaboration. “The CMS trigger system”. In: *Journal of Instrumentation* 12.01 (Jan. 2017), P01020–P01020. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/01/p01020. URL: http://dx.doi.org/10.1088/1748-0221/12/01/P01020.

[26] G L Bayatyan et al. *CMS computing: Technical Design Report*. Technical design report. CMS. Submitted on 31 May 2005. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/838359.

[27] A.M. Sirunyan et al. “Particle-flow reconstruction and global event description with the CMS detector”. In: *Journal of Instrumentation* 12.10 (Oct. 2017), P10003–P10003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/10/p10003. URL: http://dx.doi.org/10.1088/1748-0221/12/10/P10003.

[28] CMS Collaboration. *SWGuideParticleFlow (revision 51)*. Nov. 2018. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideParticleFlow?rev=51 (visited on 11/09/2021).

[29] Rene Brun and Fons Rademakers. “ROOT — An object oriented data analysis framework”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81–86. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(97)00048-X. URL: https://www.sciencedirect.com/science/article/pii/S016890029700048X.

[30] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024.

[31] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[32] Manuel Bähr et al. “Herwig++ physics and manual”. In: *The European Physical Journal C* 58.4 (Nov. 2008), pp. 639–707. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-008-0798-9. URL: http://dx.doi.org/10.1140/epjc/s10052-008-0798-9.

[33] S. Agostinelli et al. “Geant4 – a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[34] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[35] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Physical Review D* 48.7 (Oct. 1993), pp. 3160–3166. ISSN: 0556-2821. DOI: 10.1103/physrevd.48.3160. URL: http://dx.doi.org/10.1103/PhysRevD.48.3160.

[36] Yu.L Dokshitzer et al. "Better jet clustering algorithms". In: *Journal of High Energy Physics* 1997.08 (Aug. 1997), pp. 001–001. ISSN: 1029-8479. DOI: 10.1088/1126-6708/1997/08/001. URL: http://dx.doi.org/10.1088/1126-6708/1997/08/001.

[37] M Wobisch and T Wengler. "Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering". In: (July 1999).

[38] A.M. Sirunyan et al. "Pileup mitigation at CMS in 13 TeV data". In: *Journal of Instrumentation* 15.09 (Sept. 2020), P09018–P09018. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/09/p09018. URL: http://dx.doi.org/10.1088/1748-0221/15/09/P09018.

[39] A.M. Sirunyan et al. "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ using the CMS detector". In: *Journal of Instrumentation* 14.07 (July 2019), P07004–P07004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/07/p07004. URL: http://dx.doi.org/10.1088/1748-0221/14/07/P07004.

[40] V. Khachatryan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *Journal of Instrumentation* 12.02 (Feb. 2017), P02014–P02014. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/02/p02014. URL: http://dx.doi.org/10.1088/1748-0221/12/02/P02014.

[41] Giovanni Franzoni. "Dataset definition for CMS operations and physics analyses". In: *Nuclear and Particle Physics Proceedings* 273-275 (2016). 37th International Conference on High Energy Physics (ICHEP), pp. 929–933. ISSN: 2405-6014. DOI: https://doi.org/10.1016/j.nuclphysbps.2015.09.144. URL: https://www.sciencedirect.com/science/article/pii/S2405601415006331.

[42] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[43] CMS Collaboration. "Extraction and validation of a new set of CMS pythia8 tunes from underlying-event measurements". In: *The European Physical Journal C* 80.1 (Jan. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7499-4. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7499-4.

[44] A. Bodek et al. “Extracting muon momentum scale corrections for hadron collider experiments”. In: *The European Physical Journal C* 72.10 (Oct. 2012). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-012-2194-8. URL: http://dx.doi.org/10.1140/epjc/s10052-012-2194-8.

[45] CMS Collaboration. *RochcorMuon (revision 36)*. Oct. 2021. URL: https://twiki.cern.ch/twiki/bin/view/CMS/RochcorMuon?rev=36 (visited on 10/19/2021).

[46] CMS Collaboration. “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 16.05 (May 2021), P05014. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/05/p05014. URL: http://dx.doi.org/10.1088/1748-0221/16/05/P05014.

[47] CMS Collaboration. *EGamma RunII Recommendations*. URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaRunIIRecommendations?rev=19#Residual_Scale_Smearing_Correcti (visited on 10/20/2021).

[48] CMS Collaboration. “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$”. In: *Journal of Instrumentation* 13.06 (June 2018), P06015–P06015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/06/p06015. URL: http://dx.doi.org/10.1088/1748-0221/13/06/P06015.

[49] CMS Collaboration. “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 16.05 (May 2021), P05014. DOI: 10.1088/1748-0221/16/05/p05014. URL: https://doi.org/10.1088/1748-0221/16/05/p05014.

[50] “Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV”. In: (Sept. 2021). URL: https://cds.cern.ch/record/2792322.

[51] Max Fischer et al. *MatterMiners/cobald: v0.12.3*. Version v0.12.3. Nov. 2021. DOI: 10.5281/zenodo.5681401. URL: https://doi.org/10.5281/zenodo.5681401.

[52] Manuel Giffels et al. *MatterMiners/tardis: The Survivors*. Version 0.6.0. Aug. 2021. DOI: 10.5281/zenodo.5177160. URL: https://doi.org/10.5281/zenodo.5177160.

[53] Lawrence Livermore National Laboratory. *Batch System Primer | High Performance Computing*. URL: https://hpc.llnl.gov/banks-jobs/running-jobs/batch-system-primer (visited on 10/06/2021).

[54] HTCondor Team. *HTCondor*. Version 9.0.8. Dec. 2021. DOI: 10.5281/zenodo.5749916. URL: https://doi.org/10.5281/zenodo.5749916.

[55] F Pacini and A Maraschini. “Job description language (JDL) attributes specification”. In: *EGEE Consortium* 590869 (2006).

[56] I Sfiligoi. “glideinWMS—a generic pilot-based workload management system”. In: *Journal of Physics: Conference Series* 119.6 (July 2008), p. 062044. DOI: 10.1088/1742-6596/119/6/062044. URL: https://doi.org/10.1088/1742-6596/119/6/062044.

[57] Guenther Erli et al. *roced-scheduler/ROCED 1.1.0*. Version 1.1.0. Dec. 2018. DOI: 10.5281/zenodo.1888310. URL: https://doi.org/10.5281/zenodo.1888310.

[58] Fischer, Max et al. “Effective Dynamic Integration and Utilization of Heterogenous Compute Resources”. In: *EPJ Web Conf.* 245 (2020), p. 07038. DOI: 10.1051/epjconf/202024507038. URL: https://doi.org/10.1051/epjconf/202024507038.

[59] *Together We Are Powerful*. Folding@home. URL: https://foldingathome.org/ (visited on 12/08/2021).

[60] *Folding@home | Statistics*. URL: https://stats.foldingathome.org/donor/73729296 (visited on 12/08/2021).

[61] Caspart, René et al. “Setup and commissioning of a high-throughput analysis cluster”. In: *EPJ Web Conf.* 245 (2020), p. 07007. DOI: 10.1051/epjconf/202024507007. URL: https://doi.org/10.1051/epjconf/202024507007.

[62] R F von Cube et al. “Federation of compute resources available to the German CMS community”. In: *Journal of Physics: Conference Series* 1525 (Apr. 2020), p. 012055. DOI: 10.1088/1742-6596/1525/1/012055.

[63] Red Hat Inc. *Fedora*. URL: https://getfedora.org/ (visited on 10/27/2021).

[64] Chris Wright. *Transforming the development experience within CentOS*. URL: https://www.redhat.com/en/blog/transforming-development-experience-within-centos (visited on 10/06/2021).

[65] Chris Wright. *CentOS Stream: Building an innovative future for enterprise Linux*. URL: https://www.redhat.com/en/blog/centos-stream-building-innovative-future-enterprise-linux (visited on 10/06/2021).

[66] LinuxCERN. *Recent software updates and News*. URL: https://linux.web.cern.ch/#update-on-centos-linux-strategy (visited on 10/27/2021).

[67] *Open Source Cloud Computing Infrastructure*. OpenStack. URL: https://www.openstack.org/ (visited on 12/12/2021).

[68] *Apache Cloudstack*. Apache Cloudstack. URL: https://cloudstack.apache.org/ (visited on 12/12/2021).

[69] Jakob Blomer et al. “The Evolution of Global Scale Filesystems for Scientific Software Distribution”. In: *Computing in Science Engineering* 17.6 (2015), pp. 61–71. DOI: 10.1109/MCSE.2015.111.

[70] *Welcome to CernVM-FS’s documentation! — CernVM-FS 2.9.0 documentation*. URL: https://cvmfs.readthedocs.io/en/stable/ (visited on 12/12/2021).

[71] Squid. *squid: Optimising Web Delivery*. URL: http://www.squid-cache.org/ (visited on 10/06/2021).

[72] Qi Zhang et al. *A Comparative Study of Containers and Virtual Machines in Big Data Environment*. 2018. arXiv: 1807.01842 [cs.DC].

[73] Dirk Merkel. “Docker: lightweight linux containers for consistent development and deployment”. In: *Linux journal* 2014.239 (2014), p. 2.

[74] Docker Inc. *Empowering App Development for Developers*. URL: https://www.docker.com/ (visited on 10/27/2021).

[75] Gregory M. Kurtzer, Vanessa Sochat, and Michael W. Bauer. “Singularity: Scientific containers for mobility of compute”. In: *PLOS ONE* 12.5 (May 11, 2017). Ed. by Attila Gursoy, e0177459. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0177459. URL: https://dx.plos.org/10.1371/journal.pone.0177459 (visited on 10/27/2021).

[76] *Home | Singularity*. URL: https://singularity.hpcng.org/ (visited on 10/27/2021).

[77] *Production-Grade Container Orchestration*. Kubernetes. URL: https://kubernetes.io/ (visited on 10/27/2021).

[78] *What is SquashFS*. URL: https://tldp.org/HOWTO/SquashFS-HOWTO/whatis.html (visited on 10/27/2021).

[79] *NAMESPACES(7) – Linux Programmer’s Manual*. Release 5.05. Aug. 2019.

[80] B Bockelman et al. “Commissioning the HTCondor-CE for the Open Science Grid”. In: *Journal of Physics: Conference Series* 664.6 (Dec. 2015), p. 062003. DOI: 10.1088/1742-6596/664/6/062003. URL: https://doi.org/10.1088/1742-6596/664/6/062003.

[81] Puppet Webteam. *Powerful infrastructure automation and delivery | Puppet*. URL: https://www.puppet.com// (visited on 12/02/2021).

[82] Internet Engineering Task Force. *Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*. URL: https://tools.ietf.org/html/rfc5280 (visited on 10/06/2021).

[83] *CLAIX / RWTH Compute Cluster - RWTH AACHEN UNIVERSITY IT Center - Deutsch*. URL: https://www.itc.rwth-aachen.de/go/id/eucm (visited on 11/03/2021).

[84] Mark S. Birrittella et al. “Intel® Omni-path Architecture: Enabling Scalable, High Performance Fabrics”. In: *2015 IEEE 23rd Annual Symposium on High-Performance Interconnects*. 2015, pp. 1–9. DOI: 10.1109/HOTI.2015.22.

[85] *SSH(1) – BSD General Commands Manual*. Dec. 2021.

[86] *MatterMiners/container-stacks/wlcg-wn*. original-date: 2020-11-19T09:21:53Z. Oct. 29, 2021. URL: https://github.com/MatterMiners/container-stacks/blob/f4cf4cacb14217196f22cf78c7d06c4ca84660c5/wlcg-wn/Dockerfile (visited on 12/13/2021).

Danksagung

An dieser Stelle möchte ich einigen Personen meinen Dank aussprechen, die mir durch ihre Unterstützung und Motivation diese Arbeit ermöglicht haben.

Zuerst geht ein ganz besonderer Dank an Prof. Dr. Günter Quast dafür, dass er mich in die Gruppe aufgenommen und mir die Möglichkeit gegeben hat, auf diesen Gebieten zu arbeiten. Vielen Dank für die Gelegenheit, immer mit fachlichem und inbesondere motivierendem Beistand, diese themenübergreifende Arbeit verfassen zu können.

Des Weiteren möchte ich Prof. Dr. Alexander Schmidt meinen Dank aussprechen dafür, dass er dieses Projekt ermöglicht und sich auf die Aachen-Karlsruher Zusammenarbeit eingelassen, sowie sich bereit erklärt hat, das Koreferat der Arbeit zu übernehmen.

Für die gute Betreuung, den immerwährenden produktiven Austausch und das Korrekturlesen diverser Schriften, nicht zuletzt dieser Arbeit, möchte ich Dr. Manuel Giffels danken.

Dr. Andreas Nowack und Dr. Thomas Kreß danke ich, dass sie die Projekte auf Aachener Seite ermöglicht und unterstützt haben.

Für die produktive Zusammenarbeit und das oft heitere Klima in der Gruppe danke ich den vielen Personen am Institut für experimentelle Teilchenphysik. Insbesondere Dr. Daniel Săvoiu, Dr. Matthias Schnepf, Sebastian Brommer, Robin Hofsaess und Maximilian Horzela gilt ein großer Dank für Korrekturen, Diskussionen und Unterstützung in den letzten Wochen.

Zu guter Letzt möchte ich meiner Familie und meinen Freunden, die mich bei der Erstellung der Arbeit immer unterstützt und mir viel Verständnis entgegengebracht hat, von ganzem Herzen danken.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die Dissertationsschrift mit dem Titel

Dynamic Integration of
Heterogeneous Computing Resources
and Jet Energy Calibration
for the CMS Experiment

selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht habe, was aus Arbeiten Anderer unverändert oder mit Abänderung entnommen wurde.

Ich versichere außerdem, dass ich die Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

Karlsruhe, den 19.01.2022
R. Florian von Cube