


Precision calibration of calorimeter signals in the ATLAS experiment using an uncertainty-aware neural network

ATLAS collaboration



Abstract

The ATLAS experiment at the Large Hadron Collider explores the use of modern neural networks for a multi-dimensional calibration of its calorimeter signal defined by clusters of topologically connected cells (topo-clusters). The Bayesian neural network (BNN) approach not only yields a continuous and smooth calibration function that improves performance relative to the standard calibration but also provides uncertainties on the calibrated energies for each topo-cluster. The results obtained by using a trained BNN are compared to the standard local hadronic calibration and to a calibration provided by training a deep neural network. The uncertainties predicted by the BNN are interpreted in the context of a fractional contribution to the systematic uncertainties of the trained calibration. They are also compared to uncertainty predictions obtained from an alternative estimator employing repulsive ensembles.

 Copyright CERN for the benefit of the ATLAS collaboration. This work is licensed under the Creative Commons Attribution 4.0 International License. Published by the SciPost Foundation.

Received 2024-12-16

Accepted 2025-11-18

Published 2025-12-16

doi:[10.21468/SciPostPhys.19.6.155](https://doi.org/10.21468/SciPostPhys.19.6.155)



Check for updates

Contents

1	Introduction	3
2	Experimental setup and simulations	4
2.1	The ATLAS detector	5
2.2	Monte Carlo simulations	5
2.3	Calorimeter signal formation and features	6
2.3.1	Topo-cluster kinematics	6
2.3.2	Simulation data content	7
2.4	Event observables	8
3	Datasets and network setup	8
3.1	Dataset	8
3.1.1	Final-state objects and phase space	8
3.1.2	Feature selection for the calibration training	9

3.2	Topo-cluster calibration training	12
3.2.1	Training target	12
3.2.2	Feature and target transformations and projections	15
3.3	Models and network setup	16
3.3.1	Principle design of a Bayesian neural network	16
3.3.2	BNN loss function	17
3.3.3	Repulsive ensembles	18
3.3.4	Repulsive ensemble loss function	20
4	Performance evaluations	21
4.1	Evaluation metrics	21
4.1.1	Accuracy of the learned response	21
4.1.2	Signal linearity	22
4.1.3	Relative local energy resolution	22
4.2	Prediction power	23
4.2.1	Distribution shapes	23
4.2.2	Central prediction power	24
4.2.3	Network training issues	24
4.3	Central calibration performance	27
4.4	Local topo-cluster energy resolution	29
5	Analyzing the uncertainty predictions	32
5.1	Predictive uncertainties	32
5.2	Uncertainty components	33
5.3	Comparison of the learned uncertainties	33
5.4	Uncertainties and data spread	36
5.5	Sources of uncertainties	36
6	Conclusions	38
A	Local hadronic calibration sequence	41
B	Features	42
B.1	Feature construction	42
B.2	Relation of features to topo-cluster nature	43
C	Network implementations and loss functions	45
C.1	Bayesian loss function	45
C.2	Implementation of repulsive ensemble networks	46
C.3	Repulsive ensemble loss function	47
D	Evaluation metrics	48
D.1	Performance	48
D.2	Uncertainties	49
D.3	Expectations for the pull distribution	49
E	Calibration performance	50
E.1	Performance in feature space	50
E.2	Local relative energy resolution	50
	References	53

1 Introduction

The principal signals of the non-compensating calorimeters employed by the ATLAS experiment [1] at the Large Hadron Collider (LHC) are clusters of topologically connected cell signals (topo-clusters) that reconstruct locally deposited energy into a three-dimensional signal object [2]. The calorimeter signals are calibrated to correctly measure the energy deposited by electromagnetic showers, thus providing no compensation for the fraction of deposited energy not contributing to the signal in the complex development of hadronic showers. This leads to on average smaller observed signals for hadrons depositing the same energy as electrons, positrons, or photons.

In ATLAS, the standard approach to compensate for the hadronic signal deficiencies at the level of individual topo-clusters is to apply a *local hadronic calibration* as part of a full calibration chain that also includes corrections for other effects. This chain is implemented as a four-step sequence of (1) a cluster-by-cluster signal classification determining the likelihood of the topo-cluster to be of electromagnetic origin, (2) a hadronic calibration employing a local cell-signal weighting technique giving the chain its name (LCW), (3) an out-of-cluster correction, and (4) a correction for energy losses in inactive material in the proximity of the topo-cluster, as detailed in Ref. [2]. All classification probabilities and calibration and correction scale factors used in this sequence are retrieved from multi-dimensional lookup tables that are calculated from single particle response simulations. These tables provide averages in bins of observables sensitive to the nature of the energy deposit and to effects on the signal introduced by the detector geometry, and its signal formation and extraction characteristics at the topo-cluster location. In addition to bin edge effects potentially introducing discontinuities of scale factors, the correlations between the observables are reduced to average relations in this approach. This is expected to introduce limitations to the precision of the LCW calibration, as the specific relation between the observables used as inputs is lost for any given topo-cluster.

Modern machine-learning (ML) methods [3] can be used to learn multi-dimensional, continuous calibration functions without step-like transitions. These functions preserve, and even use, the underlying correlations between observables, thus promising a more accurate topo-cluster calibration at higher precision. A set of topo-cluster observables commonly referred to as *features* provides the inputs to the corresponding training, and to the inference of the trained model. This set includes physics-motivated data representations encompassing measures of the topo-cluster signal characteristics, the underlying nature of the energy deposit, and the signal environment at the LHC.

A one-step training procedure omitting the dedicated topo-cluster classification step was designed in a previous approach to determine such a function. The results obtained by training a deep neural network (DNN) to learn the response of individual topo-clusters from which the signal calibration is derived are presented in Ref. [4]. The performance of the DNN-trained calibration with respect to signal linearity and resolution was found to yield significant improvements when compared with the standard hadronic calibration provided in the LCW sequence. Other machine-learning-based approaches to calibrate topo-clusters use cells in topo-clusters in deep learning networks employing images, graphs, and point clouds. Those were successfully trained to calibrate single particles in full Monte Carlo (MC) simulations of the ATLAS detector [5, 6].

In this paper, results from the application of a Bayesian neural network (BNN) [7–13] for the calibration of the topo-clusters obtained with full MC simulations of multijet final states of the proton–proton (pp) collisions at the LHC Run 2 (2015–2018) centre-of-mass energy of $\sqrt{s} = 13$ TeV are presented. These simulations encompass physics modelling of the particle-level final state and a detailed detector simulation. The detector simulation models the signal

contributions from pile-up introduced by the high inelastic pp cross-section, in combination with the high proton beam intensities observed in Run 2.

The BNN is trained to not only predict the response and thus the calibration function cluster-by-cluster, but to also learn a statistical and a systematic uncertainty component associated with this prediction. The naming of these uncertainty components in the context of this study reflects expectations for their respective behaviour and may differ from typical interpretations in calibrations that are derived using measures of centrality in samples of signal objects, rather than directly determining them for a single object (topo-cluster) alone. Also not considered here are possible other (systematic) uncertainty contributions not directly associated with the application of the learned calibration model, including those arising from residual discrepancies between experimental data and MC simulations. An example of a complete determination of systematic uncertainties can be found in the jet calibration in ATLAS, with its approach for LHC Run 2 discussed in detail in Ref. [14].

The predicted statistical uncertainty component captures the dependence of the prediction on the size of the dataset used for training the BNN, while the systematic uncertainty component provides a measure for the accuracy of the prediction that cannot be improved by a larger training dataset or more training cycles (*epochs*). Depending on the approach to the full local hadronic calibration of topo-clusters, which, as mentioned before, potentially includes components introduced by the comparison of experimental data to MC simulations, these uncertainties can provide important contributions to the overall systematic uncertainty associated with the application of the trained BNN model in the course of such calibration.

The individual uncertainty predictions are available for all topo-clusters with a calibration learned by the BNN, including those that are not part of the signal of a calibrated physics object like jets or particles. They can be employed to improve the input signal quality by topo-cluster selections prior to the reconstruction of particles and jets. In addition to these selections, uncertainty propagations can be applied in *bottoms-up* calibration schemes used in the context of the reconstruction of more complex and softer energy flows that are potentially needed for the measurements of the missing transverse momentum, the soft hadronic recoil in vector boson production, and other hadronic event shapes, as well as in jet substructure reconstruction and analysis.

To increase confidence in the uncertainty predictions from the BNN, they are compared to an alternative uncertainty estimate determined with *repulsive ensembles* (RE) [15, 16]. Observed large predicted uncertainties for certain topo-clusters are interpreted in the context of the signal characteristics associated with complex transition regions in the ATLAS calorimeter system.

The structure of this paper is as follows. In Section 2, the ATLAS detector and the used datasets are presented. Introduction to the neural-network-based calibration, the features it uses, and the BNN and RE network architecture and configurations are given in Section 3. In Section 4, the performance of the calibration is compared to the LCW and the DNN-derived calibrations. Section 5 discusses the validity of the uncertainties predicted by the BNN in a comparison with the corresponding RE predictions and in the context of detector signal features. The paper concludes with Section 6.

2 Experimental setup and simulations

The results presented here are solely obtained with MC simulations that model the setup of the ATLAS experiment and the signal formation and extraction for the various sub-detectors in great detail. The emphasis is on the central region of the calorimeter system. It is the most relevant detector component for the studies presented in this paper, as the topo-clusters used to

learn the response and validate the obtained calibration are collected from calorimeter jets in this region. The only other detector systems contributing here are the inner tracking detector for the reconstruction of the collision vertices and the dedicated forward detectors comprising the luminosity measurement system [17] used for the reconstruction of the number of pile-up collisions.

2.1 The ATLAS detector

The ATLAS detector at the LHC covers nearly the entire solid angle around the pp collision vertex.¹ It consists of an inner tracking detector system (ID) surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer (MS) incorporating three large superconducting air-core toroidal magnets.

The calorimeter system provides near full absorption coverage in the pseudorapidity range $|\eta_{\text{det}}| < 4.9$. Electromagnetic calorimetry is provided up to $|\eta_{\text{det}}| < 3.2$ by highly granular barrel (EMB) and endcap (EMEC) lead/liquid-argon (LAr) calorimeters. An additional thin LAr presampler covers $|\eta_{\text{det}}| < 1.8$ to correct for energy losses in material in front of the calorimeters.

Hadron calorimetry up to $|\eta_{\text{det}}| < 1.7$ is provided by the steel/scintillator-tile (Tile) calorimeter with three barrel structures (one barrel Tile and two extended barrel Tile calorimeters). The two hadronic endcap (HEC) copper/LAr calorimeters provide coverage for $1.5 < |\eta_{\text{det}}| < 3.2$. Scintillating counters are installed between the barrel and the extended Tile calorimeters to measure energy losses in this transition region within $0.8 < |\eta_{\text{det}}| < 1.6$ (Tile gap scintillators).

The two forward calorimeters (FCal) cover $3.1 < |\eta_{\text{det}}| < 4.9$. Each consists of a copper/LAr module followed by two tungsten/LAr modules for a total of three longitudinal segments, a setup optimised for electromagnetic and hadronic energy measurements.

The ATLAS calorimeters feature highly granular readout, with a presampler followed by six longitudinal readout segments in the barrel region $|\eta_{\text{det}}| < 1.5$, a presampler up to $|\eta_{\text{det}}| = 1.8$ and six to seven longitudinal segments in the endcap region, and three longitudinal segments in the forward region. The calorimeter has a depth of at least 10λ everywhere, with λ being the hadronic absorption length. Overall, it has nearly 200 000 independent readout cells, with the highest lateral granularities found in the EMB and the EMEC. The software suite described in Ref. [18] is used for the reconstruction and analysis of the MC simulations. It is in large parts identical to the one used for experimental data.

2.2 Monte Carlo simulations

The fully simulated jet samples were produced according to the description given in Ref. [19]. The hard-scatter pp interactions were generated at $\sqrt{s} = 13\text{ TeV}$ using PYTHIA 8.230 [20] with the A14 [21] set of tuned parton shower and underlying event parameters and the NNPDF2.3LO [22] set of parton distribution functions (PDF).

To simulate the detector signals, the stable particles² from the generated final state were tracked through the ATLAS detector using the GEANT4 [23] toolkit. The energy deposited in the calorimeters was collected in the detector simulation and is available for each calorimeter

¹ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the centre of the LHC ring, and the y -axis points upwards. Polar coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$ and is equal to the rapidity $y = \frac{1}{2} \ln \left(\frac{E+p_z}{E-p_z} \right)$ in the relativistic limit. The pseudorapidity η_{det} in the detector frame of reference follows the definition of η with reference to the nominal collision vertex at $(x, y, z) = (0, 0, 0)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$.

²Here particles are considered stable if their proper lifetime τ is $\tau > 10\text{ ps}$.

cell and in regions of inactive material. The signal formation in the various detector systems is modelled following the respective signal extractions and digitisations in the experiment [24]. The effect of multiple pp interactions in the same (in-time pile-up) and neighbouring bunch crossings (out-of-time pile-up) was modelled by overlaying the simulated signals of the hard-scattering event with signals from inelastic *minimum bias* pp collisions generated with PYTHIA 8.186 [25] using the NNPDF2.3LO set of PDFs and the A3 set of tuned parameters [26]. The number of overlaid pile-up interactions is sampled from the distribution measured in the data during LHC Run 2.

2.3 Calorimeter signal formation and features

The formation of topo-clusters principally collects calorimeter cells starting from seed cells with highly significant signals and growing by following cell signal significance patterns within and across calorimeter modules in three dimensions. The initial clustering is followed by an algorithm splitting large topo-clusters between local signal maxima, again in three dimensions. The full procedure and the reconstructed observables associated with the final cluster are detailed in Ref. [2]. This reference also introduces the LCW calibration that uses signal characteristics, shapes and locations associated with the topo-clusters.

2.3.1 Topo-cluster kinematics

The reconstructed topo-clusters have kinematics that are represented at two scales, the basic electromagnetic (EM) scale and the calibrated LCW scale. At both scales, topo-clusters are considered to be massless *pseudo-particles*. The four-momentum representation $P_{\text{clus}}^{\text{EM}}$ of a topo-cluster at the EM scale is given by

$$P_{\text{clus}}^{\text{EM}} = E_{\text{clus}}^{\text{EM}} \cdot (1, \cos \phi_{\text{clus}}^{\text{EM}} / \cosh y_{\text{clus}}^{\text{EM}}, \sin \phi_{\text{clus}}^{\text{EM}} / \cosh y_{\text{clus}}^{\text{EM}}, \tanh y_{\text{clus}}^{\text{EM}}), \quad (1)$$

where $E_{\text{clus}}^{\text{EM}}$, $y_{\text{clus}}^{\text{EM}}$ and $\phi_{\text{clus}}^{\text{EM}}$ are the cluster energy, rapidity and azimuth reconstructed at the EM scale, respectively.

The ML calibrations derived in this study are specific for each topo-cluster and focus on the energy deposit it represents at its location in the calorimeter. There is no attempt to include predictions for the correction of energy losses outside of the topo-cluster. Such corrections are included in the full LCW chain, though, see Appendix A for more details.

The *hadronically calibrated* cluster energy $E_{\text{clus}}^{\text{had}}$ is available from the LCW sequence, before the corrections for out-of-cluster and inactive material energy losses are applied. It serves as a reference for performance comparisons with the ML calibrations. The expectation is that it represents the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the calorimeter cells collected into the topo-cluster,

$$E_{\text{clus}}^{\text{EM}} \xrightarrow{\text{hadronic calibration}} E_{\text{clus}}^{\text{had}} = \underbrace{E_{\text{clus}}^{\text{dep}}}_{\text{expectation/goal}}$$

Consequently, $E_{\text{clus}}^{\text{had}}$ is referred to as the signal at the local hadronic LCW scale, as it is reconstructed in the context of this calibration but omits further corrections not related to the energy deposit within the volume of the topo-cluster. Due to the cell signal weighting applied in the hadronic calibration step of the LCW procedure, typically $E_{\text{clus}}^{\text{had}} > E_{\text{clus}}^{\text{EM}}$, $y_{\text{clus}}^{\text{EM}} \neq y_{\text{clus}}^{\text{had}}$ and $\phi_{\text{clus}}^{\text{EM}} \neq \phi_{\text{clus}}^{\text{had}}$ are expected for all topo-clusters not classified as purely electromagnetic (EM) based on an EM-likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}} < 1$ measure provided in the classification step of this procedure. The machine-learning-based calibration applies a scale factor for $E_{\text{clus}}^{\text{EM}}$ to represent $E_{\text{clus}}^{\text{dep}}$, and as such does not change the topo-cluster direction reconstructed at EM scale.

2.3.2 Simulation data content

The detector simulation provides the relevant *truth information* that assesses the energy $E_{\text{clus}}^{\text{dep}}$ deposited inside the topo-cluster. This is the sum of all energies $E_{\text{cell}}^{\text{dep}}$ deposited in the $N_{\text{clus}}^{\text{cell}}$ clustered cells during the shower simulation of *primary particles* emerging directly from the generated hard-scatter interaction,

$$E_{\text{clus}}^{\text{dep}} = \sum_{i=1}^{N_{\text{clus}}^{\text{cell}}} w_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{dep}}.$$

The geometrical weight $0 < w_{\text{cell},i}^{\text{geo}} \leq 1$ reflects the fractional contribution of $E_{\text{cell},i}^{\text{dep}}$ to the topo-cluster if the cell i is shared with another cluster, with $w_{\text{cell},i}^{\text{geo}} = 1$ for cells i exclusively collected into one topo-cluster.

Within this approach $E_{\text{clus}}^{\text{dep}}$ is the simulated *detector truth* that is generated by the energy flow carried by primary particles into the calorimeter and collected within the constraints introduced by the topo-cluster algorithm. Thus, cells that have energy deposits from these particles, but are not clustered, do not contribute to the detector truth of a given final state. Also not included in $E_{\text{clus}}^{\text{dep}}$ are energies deposited by particles from in-time pile-up collisions, or any energies deposited by particle generated in nearby bunch crossings contributing to out-of-time pile-up. Any target for a machine-learned calibration, or any other, which is constructed from $E_{\text{clus}}^{\text{dep}}$ thus reflects only the energy deposit generated by the *pp* interaction of interest, and not from pile-up interactions. This principally allows determining calibration functions that locally, at the level of each topo-cluster, mitigate the effect of pile-up on the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$ with these simulations.

Each $E_{\text{cell}}^{\text{dep}}$, and thus by extension the $E_{\text{clus}}^{\text{dep}}$ it contributes to, comprises visible and invisible energy losses within the particle shower. The visible energy deposits typically consist of losses through ionisation and other interactions that potentially contribute to the cell signal E_{cell} , and thus the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$. Invisible losses do not have a signal contribution. They include the rest energy of stopped particles, particles absorbed or decaying without visible energy release, and energy invested in (slow) nuclear processes following an inelastic interaction between incoming shower particles and the matter they traverse. In addition, all energy invested in invisible particles like neutrinos, which can escape the detector without further trace, is included.³ The invisible losses mostly appear in the development of hadronic showers, and thus are the reason for the non-compensating signal character of the ATLAS calorimeter.

The composition of $E_{\text{cell}}^{\text{dep}}$, and therefore also $E_{\text{clus}}^{\text{dep}}$, is preserved in fractional contributions depending on the nature of the primary particles that cause the energy release into a cell volume in their interactions with the matter of the calorimeter. In particular, $E_{\text{clus}}^{\text{dep}}$ can be characterised by the fraction of energy $f_{\text{clus}}^{\text{dep,em}}$ deposited by showers from primary electrons/positrons (e^{\pm}) and photons (γ). Within the context of the studies presented in this paper, the following nomenclature is used to categorise topo-clusters according to $f_{\text{clus}}^{\text{dep,em}}$:

Electromagnetic topo-clusters

have at least 90% of $E_{\text{clus}}^{\text{dep}}$ coming from primary e^{\pm} or γ , $f_{\text{clus}}^{\text{dep,em}} > 0.9$.

³Within the ATLAS detector simulation, all energy released in the development of particle showers, and not invested into the production of secondary particles that can be further tracked through the detector setup, is collected at the location where it is produced. For the energy deposits discussed here, this is the calorimeter cell volume in which the interaction happened. The concept of stopped particles is implemented in the interaction modelling within GEANT4 for particles with energies or ranges below the corresponding configured thresholds for secondary particle tracking and thus does not necessarily reflect physical behaviour at the interaction level.

Hadronic topo-clusters

have less than 10% of $E_{\text{clus}}^{\text{dep}}$ deposited by primary e^\pm and γ , $f_{\text{clus}}^{\text{dep,em}} < 0.1$.

Composite topo-clusters

have varying fractional contributions to $E_{\text{clus}}^{\text{dep}}$ from various primary particles such that $0.1 \leq f_{\text{clus}}^{\text{dep,em}} \leq 0.9$.

This categorisation is not directly used in the training of the machine-learning-based calibrations, as it is not available for topo-clusters in the experiment. It is useful when evaluating the sensitivity of observables to the cluster signal origin. More sensitive observables are expected to better reflect the relation between $E_{\text{clus}}^{\text{dep}}$ and the signal $E_{\text{clus}}^{\text{EM}}$ it generated, which, among other considerations, suggests to include them as inputs for the calibration training, as further discussed in Section 3.1.2.

2.4 Event observables

The MC simulation samples include pile-up at levels observed in LHC Run 2, which affects the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$ and thus the calibration. Measurable indicators of the pile-up activity in a given event are therefore considered for inclusion into the training of the cluster calibration network. The two relevant observables in data and MC simulations are the number of reconstructed primary vertices N_{pV} [27] and the (measured) actual number of pile-up interactions in the event μ . Generally both N_{pV} and μ are correlated, with N_{pV} providing a direct measure for in-time pile-up activity of each bunch crossing while μ is also indicative of the out-of-time pile-up activity.⁴

3 Datasets and network setup**3.1 Dataset**

The dataset consists of selected topo-clusters reconstructed in MC simulations of full pp collision events at $\sqrt{s} = 13$ TeV with multijet final states. The events are generated and simulated as described in Section 2.2.

The topo-clusters are collected from fully reconstructed calorimeter jets in the central detector region of ATLAS, as provided by the MC simulations. Each jet establishes a calorimeter volume containing its topo-cluster constituents. With this, they not only provide a reference phase space for a group of clusters for future comparisons with experimental data, but also an environment characterised by varying topo-cluster densities covering many spatial cluster distributions also expected in data.

3.1.1 Final-state objects and phase space

Calorimeter jets are formed from topo-clusters with signals at the EM scale using the anti- k_t algorithm [28] in FASTJET [29] with a radius parameter $R = 0.4$. They are calibrated using the procedures described in Ref. [14], including pile-up corrections, the MC-simulations-derived jet energy scale (JES) calibration, and direction corrections. The considered phase space for these jets is motivated by typical analysis selections and to restrict the data used

⁴For MC simulations, μ is sampled from the distribution of $\langle\mu\rangle$, which is the number of interactions measured over a luminosity block at the LHC (duration about 30 s to 60 s) divided by the number of proton bunch crossings in this interval (average over all bunches). Due to this time-integrated construction, $\langle\mu\rangle$ provides an average measure for the out-of-time pile-up activity. For data, μ is determined from the number of pile-up interactions averaged over all bunches at a fixed position in the bunch train, again in a luminosity block [17].

for the machine-learning calibration to topo-clusters located in the high-granularity region $|\eta_{\text{det}}| < 2.5$ of the ATLAS calorimeters, where the *expressiveness* of their associated observables is expected to be the highest. It is defined using reconstructed calorimeter jet kinematics,

$$p_{T_{\text{jet}}}^{\text{JES}} > 20 \text{ GeV}, \quad \text{and} \quad |y_{\text{jet}}^{\text{JES}}| < 2.0, \quad (2)$$

where $p_{T_{\text{jet}}}^{\text{JES}}$ is the fully calibrated transverse momentum of the jet, and $y_{\text{jet}}^{\text{JES}}$ is its rapidity after all corrections.

In the context of this study particle jets in MC simulations are considered as the physics truth generating the detector truth introduced in Section 2.3.2. They are clustered from stable particles in the generated particle-level final state of the hard-scatter interaction, with the exception of muons and neutrinos, using the anti- k_t algorithm with the same radius parameter used for the calorimeter jets. Only calorimeter jets within an angular distance $\Delta R < 0.4$ to the nearest particle jet are accepted, to assure a significant amount of true signal contribution to the topo-clusters they are clustered from. No pile-up-jet tagging or removal is applied to these simulated and matched calorimeter jets.

The topo-clusters are collected from the simulated calorimeter jets passing the selection above. Given the jet definition, they are found to be approximately within $|y_{\text{clus}}^{\text{EM}} - y_{\text{jet}}^{\text{JES}}| \lesssim 0.4$, with $y_{\text{clus}}^{\text{EM}}$ being the cluster rapidity reconstructed at the EM signal scale, as introduced in Eq. 1 of Section 2.3.1. To remove clusters generated by pile-up, all clusters considered for training, testing and validation of the calibration networks are required to have a true energy content of $E_{\text{clus}}^{\text{dep}} > 300 \text{ MeV}$. Most topo-clusters not satisfying this selection were found to have a high level of randomness in their signal and other associated observables, with only a weak or no relation to $E_{\text{clus}}^{\text{dep}}$ at all.⁵ They are excluded as they are not significantly contributing to the hadronic final state (and jet) signals in the context of this study.

The collected sample of topo-clusters is randomised and all information from the jet context is dropped, as it is not relevant for training the networks and the performance evaluation of the learned calibration.

3.1.2 Feature selection for the calibration training

In addition to the kinematics represented by $E_{\text{clus}}^{\text{EM}}$ and $y_{\text{clus}}^{\text{EM}}$, each topo-cluster is associated with a set of observables that are canonically referred to as *cluster moments*. These are part of the general analysis data model in ATLAS. The moments were designed to provide signal characteristics including its timing and significance as well as selected measures indicative of the nature of the energy deposit in the topo-cluster, together with other measures that have impact on a cluster classification by signal sources and a local hadronic calibration.

The topo-cluster moments that are expected to be useful for a local calibration are those that establish an observable space that represents the relation between the signal $E_{\text{clus}}^{\text{EM}}$ and its source $E_{\text{clus}}^{\text{dep}}$. This relation can only be directly established, within the limitations of the models applied, in MC simulations. When using a selection of these observables for a machine-learning-based multi-dimensional topo-cluster calibration, not only the individual observables but also the correlations between them are important.

In machine-learning, the inputs to a classification or regression network are often referred to as features. For learning the calibration, each topo-cluster is represented by a *feature set* $\mathcal{X}_{\text{clus}}$ that serves as input to a learned calibration network. There are several important aspects to be considered when collecting some of the features summarised in Table 1 into $\mathcal{X}_{\text{clus}}$. These mostly relay to

⁵Per ATLAS convention, and following the discussion in Section 2.3.2, the topo-clusters generated purely by pile-up alone have $E_{\text{clus}}^{\text{dep}} = 0$. A meaningful target for calibrations requires $E_{\text{clus}}^{\text{dep}} > 0$, at least.

1. the expectation for the achievable accuracy of the calibration not only for the MC simulations but also for the application to experimental data;
2. the expressiveness of each feature with respect to the calorimeter signal source and its reflection in the signal definition given by the topological cell clustering;
3. and a meaningful feature reconstruction with sufficient resolution for almost all topo-clusters, if not all, reconstructed to measure the final state of the pp collisions in the ATLAS calorimeters.

Both aspects captured in item 1 and item 2, respectively, are tightly connected. Features with little sensitivity to the topo-cluster signal source are irrelevant as inputs and are thus omitted. Similarly, the selection of expressive but highly correlated features may lead to an unnecessarily large numerical system decreasing the computing performance and potentially introducing a higher level of *stochasticity* (noise).

The application of the topo-cluster calibration learned from MC simulations to experimental data requires that the individual features and their correlations are well modelled. This is generally achievable for individual features that are lower order topo-cluster moments, as documented in Ref. [2]. Several highly expressive second-order moments are not considered, as they are too sensitive to the (hadronic) shower modelling and the signal calculation in the simulation including pile-up. In particular, the modelling of pile-up contributions to these complex observables requires minimum bias physics and detector simulations to be accurate at often small scales that, at the present state-of-art, have not been achieved in all aspects. The second order moments that provide measurements of the longitudinal and lateral topo-cluster signal dispersion ($\langle m_{\text{long}}^2 \rangle$, $\langle m_{\text{lat}}^2 \rangle$) are included in $\mathcal{X}_{\text{clus}}$, because they were found to be largely insensitive to modelling issues and pile-up [2]. In addition, the signal compactness measure $p_T D$ and the energy-squared-weighted variance of the cell time distribution $\text{Var}_{\text{clus}}(t_{\text{cell}})$ inside the topo-cluster, both calculated as described in Appendix B.1, are included because by construction they are expected to have little sensitivity to potential (small scale) modelling issues.

The modelling of the feature correlations can only be evaluated in the context of a full performance study in which the trained calibration is applied to data, with well defined metrics. This study is out-of-scope for this method paper.

The final important aspect, noted in item 3 in the list above, must be considered when composing $\mathcal{X}_{\text{clus}}$. All chosen features need to be reconstructed for each cluster in a large sample of topo-clusters in a meaningful way, with the highest possible resolution. This requires least impact from the topo-cluster formation algorithm and cell-level detector inefficiencies due to, e.g., insufficient granularity and noise, on their expressiveness for the calibration. For this reason, the jet phase space, and consequently the topo-cluster phase space, covered by this proof-of-principle study of the applicability of a ML calibration and the understanding of its predictions, is restricted to the ATLAS calorimeter region with the highest readout granularity, as mentioned in Section 3.1.1.

Even within the chosen phase space, very few topo-clusters in the used dataset may still not have enough cells or suffer from other deficiencies introduced by the detector or the collision environment such that it is not possible to reconstruct some or all of the features with the highest possible accuracy and precision. Nevertheless, and mostly to avoid selection biases and thus predict the most universal calibration function for all topo-clusters in the illuminated calorimeter region, these clusters are represented as well in both training and testing of the calibration network.

Table 1: The topo-cluster features used in the training of the topo-cluster calibration networks, as summarised in Eq. 3. The variables are defined in Ref. [2], except for $p_T D$ and $\text{Var}_{\text{clus}}(t_{\text{cell}})$ defined in Appendix B. Selected features used in the LCW calibration approach are indicated.

Category	Symbol	LCW	Comment
Kinematics	$E_{\text{clus}}^{\text{EM}}$	yes	Signal at the electromagnetic energy scale
	$y_{\text{clus}}^{\text{EM}}$	yes	Rapidity at the electromagnetic energy scale
Signal strength, time structure	$\zeta_{\text{clus}}^{\text{EM}}$	no	Signal significance
	$\text{Var}_{\text{clus}}(t_{\text{cell}})$	no	Variance of t_{cell} distribution
Shower depth, shower shape, compactness	λ_{clus}	yes	Distance of the centre-of-gravity from the calorimeter front face (along principal cluster axis)
	$ \vec{c}_{\text{clus}} $	no	Distance of the centre-of-gravity from the nominal vertex
	f_{emc}	no	Fraction of energy in the electromagnetic calorimeter
	$\langle \rho_{\text{cell}} \rangle$	yes	Cluster signal density measure
	$\langle m_{\text{long}}^2 \rangle$	no	Normalised energy dispersion along the principal cluster axis
	$\langle m_{\text{lat}}^2 \rangle$	no	Normalised energy dispersion perpendicular to the principal cluster axis
	$p_T D$	no	Signal compactness measure (inspired by Ref. [30])
Topology	f_{iso}	yes	Cluster isolation measure
Pile-up	t_{clus}	no	Signal timing
	N_{PV}	no	Number of reconstructed primary vertices
	μ	no	Number of pile-up interactions per bunch crossing

The final composition of $\mathcal{X}_{\text{clus}}$ contains the 15 features described in Table 1,

$$\mathcal{X}_{\text{clus}} = \underbrace{\{E_{\text{clus}}^{\text{EM}}, y_{\text{clus}}^{\text{EM}}\}}_{\text{kinematics}} \underbrace{\{\zeta_{\text{clus}}^{\text{EM}}, \text{Var}_{\text{clus}}(t_{\text{cell}}), \lambda_{\text{clus}}, |\vec{c}_{\text{clus}}|, \langle \rho_{\text{cell}} \rangle, \langle m_{\text{long}}^2 \rangle, \langle m_{\text{lat}}^2 \rangle, p_T D, f_{\text{emc}}\}}_{\text{shower nature (position, compactness, signal density and internal time structure)}} \underbrace{\{f_{\text{iso}}, t_{\text{clus}}, N_{\text{PV}}, \mu\}}_{\text{pile-up}} \underbrace{\{f_{\text{iso}}\}}_{\text{topology (isolation)}}. \quad (3)$$

It reflects the guidance discussed above.⁶ Included are the kinematic variables $E_{\text{clus}}^{\text{EM}}$ and $y_{\text{clus}}^{\text{EM}}$, both basic inputs for any calorimeter calibration. Also included are all of the topo-cluster observables employed by the classification and the hadronic calibration in the LCW sequence, thus removing the need for an explicit cluster classification but preserving the information entering this step. In addition, the event-level pile-up measures N_{PV} and μ and the topo-cluster isolation measure f_{iso} are included to represent topological (varying cluster densities

⁶The DNN trained in the context of the studies presented in Ref. [4] employs the same feature set.

in jets) and environmental (pile-up) effects on the cluster signal. All members of $\mathcal{X}_{\text{clus}}$ are available for both experimental data and MC simulations.

3.2 Topo-cluster calibration training

The hadronic calibration in the LCW sequence aims at the (true) deposited energy $E_{\text{clus}}^{\text{dep}}$ as a calibration target. The determination of the local topo-cluster calibration using machine learning techniques introduces a set of novel aspects. Rather than learning $E_{\text{clus}}^{\text{dep}}$ with a typical value ranging from $\mathcal{O}(100\text{MeV})$ to $\mathcal{O}(1\text{TeV})$ directly, the response $\mathcal{R}_{\text{clus}}^{\text{EM}}$ as a function of the feature set $\mathcal{X}_{\text{clus}}$ from Eq. 3 is learned for each topo-cluster in a multi-dimensional regression fit. The target $\mathcal{R}_{\text{clus}}^{\text{EM}}$ is defined as the ratio of the cluster signal $E_{\text{clus}}^{\text{EM}}$ at EM scale and the deposited energy in the cluster $E_{\text{clus}}^{\text{dep}}$,

$$\mathcal{R}_{\text{clus}}^{\text{EM}} = E_{\text{clus}}^{\text{EM}}/E_{\text{clus}}^{\text{dep}}. \quad (4)$$

It is available in MC simulations for each topo-cluster extracted from the jet and selected according to the criteria introduced in Section 3.1.1. The significantly reduced value range for $\mathcal{R}_{\text{clus}}^{\text{EM}}$ lowers the computational effort in the training of the network by reducing the number of epochs, with the promise of a more accurate and precise calibration.

Both the BNN and the RE learn uncertainties related to the trained calibration for each topo-cluster individually. These uncertainties allow for a better understanding of (1) detector signal features, (2) possible signal quality issues in the data, and (3) possible fundamental limitations and other issues associated with the network training. As further discussed in Section 5.1, they are learned for each topo-cluster to provide a measure of the accuracy achieved with the trained calibration network.

3.2.1 Training target

In addition to using the topo-cluster response $\mathcal{R}_{\text{clus}}^{\text{EM}}$ as a calibration target, instead of the deposited energy $E_{\text{clus}}^{\text{dep}}$ that is the direct target of the LCW hadronic calibration, the dedicated classification step characteristic for the LCW procedure is omitted when training the BNN and the RE network. Instead, the features $E_{\text{clus}}^{\text{EM}}$, $y_{\text{clus}}^{\text{EM}}$, λ_{clus} and $\langle\rho_{\text{cell}}\rangle \equiv \rho_{\text{clus}}$ used in the LCW classification are included, as presented in Table 1 and Eq. 3. The result of the BNN(RE) training is the individual response prediction $\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})$ for a given topo-cluster with a feature set $\mathcal{X}_{\text{clus}}$, with

$$\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}}) \stackrel{\text{trained}}{=} \mathcal{R}_{\text{clus}}^{\text{EM}}(\mathcal{X}_{\text{clus}}), \quad (5)$$

with $\mathcal{R}_{\text{clus}}^{\text{EM}}$ from Eq. 4. Both networks are independently trained to encode the respective representation $\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}$ of the target $\mathcal{R}_{\text{clus}}^{\text{EM}}$ as a function of $\mathcal{X}_{\text{clus}}$ given in Eq. 3. The derivation of the corresponding calibrated topo-cluster energy $E_{\text{clus}}^{\text{BNN(RE)}}$ then employs $\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}$ such that

$$E_{\text{clus}}^{\text{BNN(RE)}} = \frac{E_{\text{clus}}^{\text{EM}}}{\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})}. \quad (6)$$

Figure 1 shows the complex distributions for $E_{\text{clus}}^{\text{dep}}$, $E_{\text{clus}}^{\text{EM}}$ and $\mathcal{R}_{\text{clus}}^{\text{EM}}$ for the three topo-cluster categories defined in Section 2.3.2. The particular distribution shapes observed for both $E_{\text{clus}}^{\text{dep}}$ and $E_{\text{clus}}^{\text{EM}}$ are driven by the multijet final state simulation configuration that determines the incoming flow of particles from the selected jets in the MC simulations, with its calorimeter representation subjected to the topo-cluster selection discussed in Section 3.1.1.

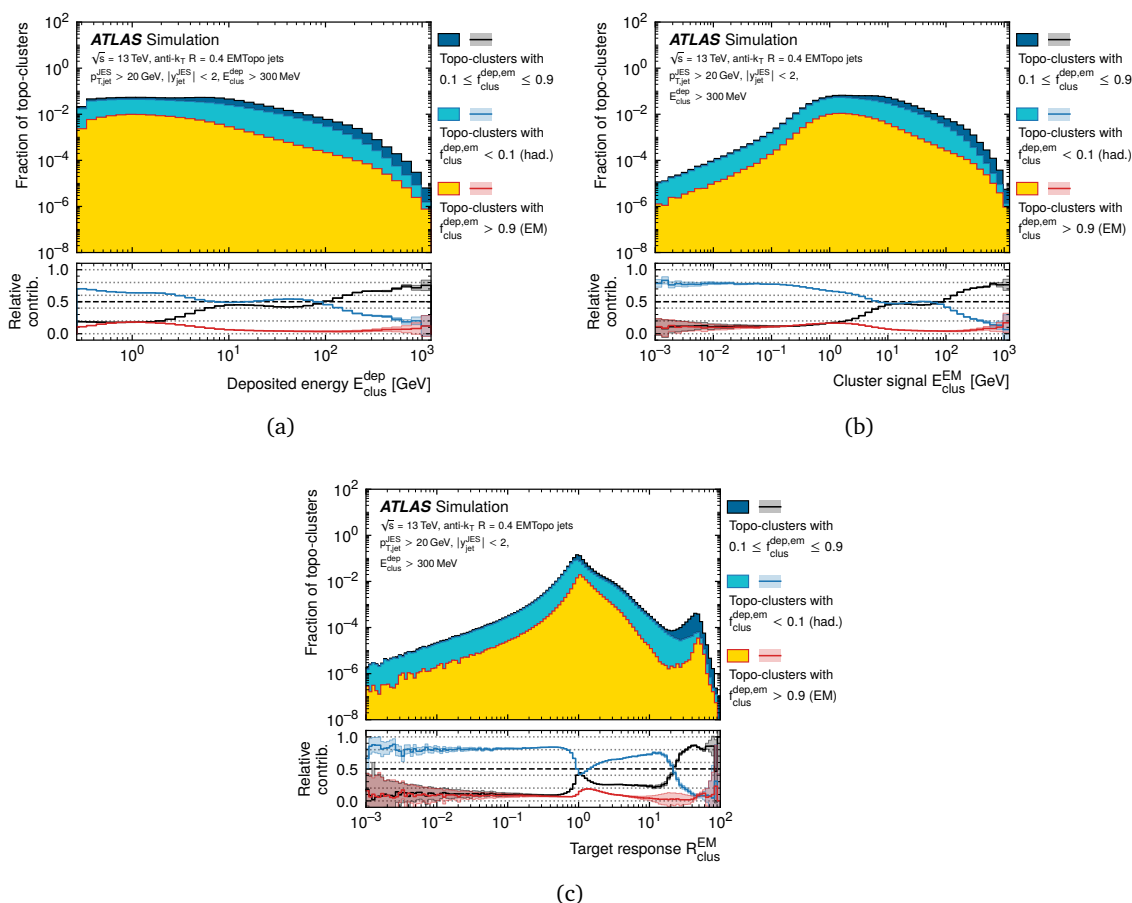


Figure 1: The distributions of the energy $E_{\text{clus}}^{\text{dep}}$ deposited in topo-clusters are shown in (a) in a stack, for the clusters categorised as electromagnetic, hadronic or composite by applying the classifications introduced in Section 2.3.2. In (b), the corresponding stacked distributions for the basic cluster signal $E_{\text{clus}}^{\text{EM}}$ at EM scale are shown. The distributions of the machine-learning training target, the topo-cluster response $\mathcal{R}_{\text{clus}}^{\text{EM}}$ constructed from $E_{\text{clus}}^{\text{EM}}$ and $E_{\text{clus}}^{\text{dep}}$ according to Eq. 4, are shown stacked in (c) for the same three categories. The distributions are filled from the sample of topo-clusters collected from fully simulated calorimeter jets, as described in Section 3.1.1. The respective lower panels show the relative contributions from the three categories to the inclusive spectrum. The shaded areas indicated the statistical errors.

The distributions provide an integral view on regional hadronic signal characteristics, like the varying levels of non-compensation across the illuminated topo-cluster rapidity range.

The contribution of electromagnetic topo-clusters to the $E_{\text{clus}}^{\text{EM}}$ distribution in Figure 1(b) is near constant for $E_{\text{clus}}^{\text{EM}} \lesssim 2 \text{ GeV}$. The lower $E_{\text{clus}}^{\text{EM}}$ regime is largely populated by hadronic topo-clusters, with their relative contribution dropping to match the one from composite topo-clusters for $10 \text{ GeV} \lesssim E_{\text{clus}}^{\text{EM}} \lesssim 100 \text{ GeV}$. Composite topo-clusters then dominate in the high signal regime $E_{\text{clus}}^{\text{EM}} > 100 \text{ GeV}$, indicating a dominant contribution from clusters generated by merged showers and shower fragments.

The resulting distributions of $\mathcal{R}_{\text{clus}}^{\text{EM}}$ shown on Figure 1(c) for the topo-cluster categories depict the complexity of the calibration task. Small $\mathcal{R}_{\text{clus}}^{\text{EM}} \ll 1$ indicate that the contribution of the deposited energy $E_{\text{clus}}^{\text{dep}}$ to the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$ is insignificant. This is the case for $E_{\text{clus}}^{\text{dep}}$ predominantly arising from stopped particles with invisible decay products like neutrinos together with a small fraction of $E_{\text{clus}}^{\text{dep}}$ generating the signal, or in deep inelastic hadronic interactions where a large amount of invisible energy is produced. In addition, the shower-generated signal reflecting $E_{\text{clus}}^{\text{dep}}$ may be suppressed because the total signal $E_{\text{clus}}^{\text{EM}}$ of the topo-cluster is dominated by negative cell signals from out-of-time pile-up noise [2].

Many of the topo-clusters with large responses $\mathcal{R}_{\text{clus}}^{\text{EM}} \gg 1$ have a centre-of-gravity located in the Tile gap scintillators [31], where the response reflects detector signal characteristics that are not only related to pile-up. These counters are not calorimeters in that they have no internal absorber structures generating showers. Their signal $E_{\text{clus}}^{\text{EM}}$ at the EM scale is calibrated to (on average) correct for energy losses in the inactive materials around them. Consequently, with the definition of $\mathcal{R}_{\text{clus}}^{\text{EM}}$ given in Eq. 4, the deposited energy $E_{\text{clus}}^{\text{dep}}$ being small in the absence of an absorber, and the signal $E_{\text{clus}}^{\text{EM}}$ by design representing energies much larger than $E_{\text{clus}}^{\text{dep}}$, a large $\mathcal{R}_{\text{clus}}^{\text{EM}}$ is expected.

The energy loss of particles produced in in-time pile-up interactions, or the secondary particles they generate by showering in the material around the Tile gap scintillators, is not accounted for in $E_{\text{clus}}^{\text{dep}}$ that is conceptionally defined in Section 2.3.2 as the detector truth generated by the final state of the hard scatter pp interaction. Due to the topo-cluster selection presented in Section 3.1.1, each cluster has a finite part of its signal $E_{\text{clus}}^{\text{EM}}$ generated by $E_{\text{clus}}^{\text{dep}} > 300 \text{ MeV}$, with the remainder generated by a contribution from pile-up that is subject to considerable fluctuations. The peak of the $\mathcal{R}_{\text{clus}}^{\text{EM}}$ distribution around $\mathcal{R}_{\text{clus}}^{\text{EM}} \approx 50$ manifests these characteristic signal features. It hints at a direct proportionality between $E_{\text{clus}}^{\text{EM}}$ and $E_{\text{clus}}^{\text{dep}}$ for the bulk of the topo-clusters population under it, which is expected for the signal from ionisation-only energy deposits in a scintillating counter.⁷

Fluctuating levels of pile-up, in terms of energy flow density and particle types in both prompt emissions from the interactions in the proton bunch crossing and in the shower development in the materials in proximity of the Tile gap scintillators caused by those, are dominant contributions to the width and shape of this peak, aside from possible modulations introduced by the topo-cluster selection applied in this study. These pile-up-induced response fluctuations arise from the pp bunch-crossing-to-bunch-crossing variations of the pile-up activity, in combination with the amount of material located between the Tile gap scintillators and the collision vertex region that comprises calorimeter modules as well as non-instrumented

⁷The energy $E_{\text{clus}}^{\text{dep}}$ in a Tile gap scintillator only represents the sum of all ionisation losses from charged particles emerging from the hard scatter interaction and passing through this detector. With the ionisation loss in the thin scintillator being about the same for each of those particles, within narrow Landau distributions [32, 33], $E_{\text{clus}}^{\text{dep}}$ is expected to be proportional to the number of charged particles passing through it. This charged particle multiplicity in turn is an indirect measure of energy losses in inelastic interactions occurring in the inactive material located around the scintillator. Thus, $E_{\text{clus}}^{\text{dep}}$ is approximately proportional to these losses when induced by particles from the hard scatter final state, but does not capture energy losses induced by charged particles directly emerging from in-time pile-up interactions or generated in showers generated by these particles.

mechanical structures. This upstream material acts as shielding that, when folded with the particle flow from the vertex region contributing to the reconstructed jets used for this study, introduces fluctuations of $E_{\text{clus}}^{\text{EM}}$ (and thus $\mathcal{R}_{\text{clus}}^{\text{EM}}$) of varying magnitude for topo-clusters between the lower rapidity edge of the Tile gap scintillator coverage at $|y_{\text{clus}}^{\text{EM}}| \approx 0.8$ (largest amount of upstream material) and the higher rapidity edge at $|y_{\text{clus}}^{\text{EM}}| \approx 1.8$ (lowest amount of upstream material). Some of these observations are revisited in the context of the predicted uncertainties in Section 5.5.

Larger $\mathcal{R}_{\text{clus}}^{\text{EM}}$ are also observed for some topo-clusters located in the electromagnetic calorimeters, which are closest to the collision region and thus exposed to the highest flow of particles from pile-up collisions. Generally, and in extension to the interpretations of the $E_{\text{clus}}^{\text{dep}}$ and $E_{\text{clus}}^{\text{EM}}$ spectra given above, the complexity of the $\mathcal{R}_{\text{clus}}^{\text{EM}}$ distributions reflect the incoming particle flow, the collision environment, and the varying detector geometries and, consequently, the varying calorimeter signal characteristics.

3.2.2 Feature and target transformations and projections

To limit the numerical value range of the network inputs and to create more appropriate (peaked) feature distributions for the training whenever possible, the features are transformed following three different procedures. If the spectrum of a feature $x \in \mathcal{X}_{\text{clus}}$ is smoothly falling and $x > 0$ always, x is replaced by a *standardised logarithmic representation*,

$$x \rightarrow \frac{\log_{10} x - \langle \log_{10} x \rangle}{\text{std}(\log_{10} x)}. \quad (7)$$

Here $\langle \log_{10} x \rangle$ is the statistical average of the $\log_{10} x$ distribution from the training sample, and $\text{std}(\log_{10} x)$ denotes the corresponding standard deviation of the transformed features x .

For other distribution shapes, x is standardised without additional transformation (*linear standardisation*),

$$x \rightarrow \frac{x - \langle x \rangle}{\text{std } x}, \quad (8)$$

where $\langle x \rangle$ and $\text{std } x$ are the mean and the standard deviation of the distribution of x from the training sample. An exception is the cluster time t_{clus} , which is normalised to its maximum absolute value $\max(|t_{\text{clus}}|)$ found in the training sample, to keep its sparsity (*maximum-absolute normalisation*),

$$t_{\text{clus}} \rightarrow \frac{t_{\text{clus}}}{\max(|t_{\text{clus}}|)}. \quad (9)$$

Table 2: Summary of the transformations applied to the various observables collected into the feature set $\mathcal{X}_{\text{clus}}$ from Eq. 3, and the target response $\mathcal{R}_{\text{clus}}^{\text{EM}}$ from Eq. 4.

Transformation (preprocessing)	Features
Logarithmic (\log_{10}) standardisation (Eq. 7)	$E_{\text{clus}}^{\text{EM}}, \zeta_{\text{clus}}^{\text{EM}}, \text{Var}_{\text{clus}}(t_{\text{cell}}), \lambda_{\text{clus}}, \langle \rho_{\text{cell}} \rangle$
Linear standardisation (Eq. 8)	$y_{\text{clus}}^{\text{EM}}, \vec{c}_{\text{clus}} , f_{\text{emc}}, \langle m_{\text{long}}^2 \rangle, \langle m_{\text{lat}}^2 \rangle, p_{\text{T}}D, f_{\text{iso}}, N_{\text{PV}}, \mu$
Maximum-absolute normalisation (Eq. 9)	t_{clus}
Logarithmic (\log_{10}) transformation (Eq. 10)	$\mathcal{R}_{\text{clus}}^{\text{EM}}$

Table 3: The architecture and hyper-parameters for both the BNN and the RE network with the three-mode Gaussian mixture likelihood. Both networks feature a decaying learning rate (LR) implemented by the learning-rate scheduler STEPLR from the PYTORCH [34] software tool kit, with LR reduced by a factor γ at each of the epochs within the range specified in braces below.

Hyper-parameter	Network architecture and setup
Likelihood model	Gaussian mixture
Number of mixture modes N_{mix}	3
Number of hidden layers	4
Number of nodes per layer (input, hidden, and output)	{15, 64, 64, 64, 64, 9}
Activation function (inner layers)	rectified linear unit (ReLU)
Activation function (last layer)	none
Central-value prediction (mean or maximum)	maximum of the likelihood
Optimizer and learning rate (LR)	ADAM [35] with LR = 10^{-4}
Learning-rate scheduler	STEPLR, epochs {25, 100}, $\gamma = 0.1$
Number of training epochs	150
Training (and inference) batch size	4096 (512)
Dataset size (all selected topo-clusters)	14.5×10^6
Data sample size for training (N_{train})	8.7×10^6 (60.0%)
Data sample size for validation	5.0×10^5 (3.4%)
Data sample size for testing	5.3×10^6 (36.6%)
Sampling at inference on the test dataset (BNN only, N)	50 times
Number of ensemble members (RE only, N)	50

For the regression target, a simple logarithmic transform is applied to reduce its numerical range further,

$$\mathcal{R}_{\text{clus}}^{\text{EM}} \rightarrow \log_{10} \mathcal{R}_{\text{clus}}^{\text{EM}}. \quad (10)$$

The transformations of the individual features listed in Table 1 and the regression target from Eq. 4 are summarised in Table 2.

3.3 Models and network setup

To learn the response given in Eq. 4 as a function of $\mathcal{X}_{\text{clus}}$, both a Bayesian neural network and a repulsive ensemble are employed. Both are configured with the same network model, with the hyper-parameters presented in Table 3. They differ significantly in the implementation of the variational approach to determine the uncertainties, and thus in the set of learned network parameters and the respective loss functions, as further discussed in Sections 3.3.2 and 3.3.4.

3.3.1 Principle design of a Bayesian neural network

As previously mentioned, Bayesian neural networks do not only predict the target of regressions like the one exercised in the context of the topo-cluster calibration, but also provide uncertainty predictions associated with the central value prediction. Figure 2 illustrates the principle setup for the training of the BNN and its inference on the (test) data.

The most significant deviation of the BNN model from the usual deep neural network design is that the weights connecting the nodes in adjacent layers are not trained as fixed values, rather they are represented by learned average values and widths of probability distributions $q(\theta)$.⁸ Here θ are the trained network parameters. When the BNN is inferred on test data, these distributions are sampled a fixed number of times, denoted by N and set to $N = 50$ for this result. This is the same as collecting the predictions from N different networks trained using the same data, without the need to do this training N times. The central prediction $\mathcal{R}_{\text{clus}}^{\text{BNN}}$ is then constructed from the N sampled predictions such that it reflects the prediction with the highest probability to represent the target $\mathcal{R}_{\text{clus}}^{\text{EM}}$, as indicated in Figure 2. The uncertainties are constructed as indicated in the same figure. They are further discussed in Section 5.

3.3.2 BNN loss function

The loss function employed in the BNN training is designed to maximise the likelihood $p(\theta|D_{\text{train}})$ of the network parameters θ to describe the training dataset D_{train} . This dataset is represented by a vector of target and feature set pairs $\{\mathcal{R}_{\text{clus}}^{\text{EM}}, \mathcal{X}_{\text{clus}}\}_j$ for $j = 1, \dots, N_{\text{train}}$ topo-clusters,

$$D_{\text{train}} = \left(\{\mathcal{R}_{\text{clus}}^{\text{EM}}, \mathcal{X}_{\text{clus}}\}_1, \{\mathcal{R}_{\text{clus}}^{\text{EM}}, \mathcal{X}_{\text{clus}}\}_2, \dots, \{\mathcal{R}_{\text{clus}}^{\text{EM}}, \mathcal{X}_{\text{clus}}\}_{N_{\text{train}}} \right). \quad (11)$$

The form of $p(\theta|D_{\text{train}})$ is not known a priori. Therefore, and following Bayes' theorem, the posterior $p(D_{\text{train}}|\theta)$ that provides the likelihood that D_{train} describes θ , is maximised instead of $p(\theta|D_{\text{train}})$, as discussed in Ref. [3]. Assuming a Gaussian likelihood for $p(D_{\text{train}}|\theta)$ is well motivated, but in case of the topo-cluster-response predictions, a mixture of $N_{\text{mix}} = 3$ Gaussian functions is chosen to accommodate possible deviations from this shape,

$$p(D_{\text{train}}|\theta) = p(\{\mathcal{R}_{\text{clus}}^{\text{EM}}\}|\theta, \{\mathcal{X}_{\text{clus}}\}) = \prod_{j=1}^{N_{\text{train}}} \sum_{i=1}^{N_{\text{mix}}} \alpha_{\theta,i} \mathcal{N}(\mathcal{R}_{\text{clus},j}^{\text{EM}} | \langle \mathcal{R} \rangle_{\theta,i}, \sigma_{\theta,i}). \quad (12)$$

Here $\alpha_{\theta,i}$ are the weights with which the Gaussian functions $\mathcal{N}(\mathcal{R}_{\text{clus}}^{\text{EM}} | \langle \mathcal{R} \rangle_{\theta,i}, \sigma_{\theta,i})$ with mean $\langle \mathcal{R} \rangle_{\theta,i}$ and standard deviation $\sigma_{\theta,i}$ contribute to $p(D_{\text{train}}|\theta)$, with

$$\alpha_{\theta,i} \geq 0, \quad \text{and} \quad \sum_{i=1}^{N_{\text{mix}}} \alpha_{\theta,i} = 1.$$

All $\alpha_{\theta,i}$, $\langle \mathcal{R} \rangle_{\theta,i}$ and $\sigma_{\theta,i}$ are predicted by the BNN and are thus functions of $\mathcal{X}_{\text{clus}}$. The mixture weights $\alpha_{\theta,i}$ are forced to be positive and to add up to unity by applying a *softmax* [36, 37] function.

To find the maximum of $p(D_{\text{train}}|\theta)$, the negative logarithmic likelihood $-\log p(D_{\text{train}}|\theta)$ is minimised. As the uncertainties $\sigma_{\theta,i}$ are learned as well, rather than being fixed *a-priori*, this approach implements a *heteroscedastic* loss function that regularises the otherwise prevalent trend of the minimisation to prefer larger values of $\sigma_{\theta,i}$ in the exponents of the Gaussian functions. Therefore, trivial solutions $\sigma_{\theta,i} \rightarrow \infty$ are avoided and the values of $\sigma_{\theta,i}$ are generally restricted to meaningful ranges. The introduction of a (reverse) Kullback-Leibler divergence $D_{\text{KL}}[q(\theta), p_{\text{prior}}(\theta)]$ [38] in the BNN loss function \mathcal{L}_{BNN} acts as a ridge (L2) regularisation [39] that controls over-fitting and focuses the fit on regions of relevant data,

$$\mathcal{L}_{\text{BNN}} = D_{\text{KL}}[q(\theta), p_{\text{prior}}(\theta)] - \underbrace{\langle \log p(D_{\text{train}}|\theta) \rangle}_{\substack{\theta \text{ sampled} \\ \text{from } q(\theta)}} \quad (13)$$

⁸The distributions $q(\theta)$ are chosen to be Gaussian functions. This is not required but enhances computational performance without observable effects on the accuracy of the predictions [11].

The $D_{\text{KL}}[q(\theta), p_{\text{prior}}(\theta)]$ term vanishes when the distribution of learned network parameters $q(\theta)$ matches the prior $p_{\text{prior}}(\theta)$. Both $q(\theta)$ and $p_{\text{prior}}(\theta)$ are chosen to be Gaussians so that $D_{\text{KL}}[q(\theta), p_{\text{prior}}(\theta)]$ can be calculated analytically. The prior mean and standard deviation are set to zero and one, respectively. The likelihood $p(D_{\text{train}}|\theta)$ is taken from Eq. 12 and the network parameters θ are sampled once from $q(\theta)$ to calculate the average negative logarithmic likelihood term in Eq. 13. More details of the motivation and the design of \mathcal{L}_{BNN} are given in Appendix C.1.

Compared to a regression training without uncertainty predictions, the network output in case of a loss defined by a single Gaussian function, configured by using $N_{\text{mix}} = 1$ in Eq. 12, is doubled. It consists of a local central value function as a representation of the regression target and a local uncertainty map. Generally, the number of network outputs increases by a factor of $3N_{\text{mix}} - 1$ for a mixture of N_{mix} Gaussian distributions.

The predictions from the BNN are a novel approach to learn uncertainty contributions to calorimeter signal calibrations. They can represent relevant contributions to the nuisances characterising the overall systematic uncertainty associated with the application of the learned local (topo-cluster by topo-cluster) calibration in data and MC simulations. The learned uncertainties also serve as indicators of the detector signal expressiveness in terms of the measurement of the true energy deposit, as shown in Section 5.5. They can thus be used as a base for topo-cluster selections to improve the accuracy and precision of the energy measurement for reconstructed particles inside and outside of jets, and thus the general energy flow in the final state of the pp collisions at the LHC.

3.3.3 Repulsive ensembles

To gain confidence in the BNN uncertainty predictions for the local topo-cluster calibration, an alternative and independent model for learning such uncertainties employing a network of repulsive ensembles is considered. The basic idea of RE is to determine uncertainties for the learned parameters θ , and consequently the functions they describe, by forcing ensembles of identically configured training networks to not predict the same best-fit network parameters, but rather explore the landscape around the corresponding minimum of the loss. While some variation of the predictions between members of an ensemble of networks can be observed due to random network initialisation and random training batch selection, there is no explicit mechanism that guarantees that the full weight space encoding $p(D_{\text{train}}|\theta)$ is mapped out. It can be shown that adding a repulsive force (network interaction) to the usual update rule that minimises $-\log p(D_{\text{train}}|\theta)$ based on gradient descent, maps the weight space completely when applied to all members of the network ensemble simultaneously at each update $t \rightarrow t + 1$ evolving $\theta^t \rightarrow \theta^{t+1}$. This procedure allows predicting uncertainties, see, e.g., the comprehensive discussion in Ref. [3].

The basic design of the RE training model is sketched in Figure 3. While the formal determination of the central prediction and the uncertainties shown in this figure is identical to the one for the BNN shown in Figure 2, the underlying sampling strategies are fundamentally different. The BNN directly samples in weight space, while the RE samples in the function space establishing the repulsive force between networks. Some relevant details of the complex theoretical foundation and the implementation of RE are given in Appendix C.2.

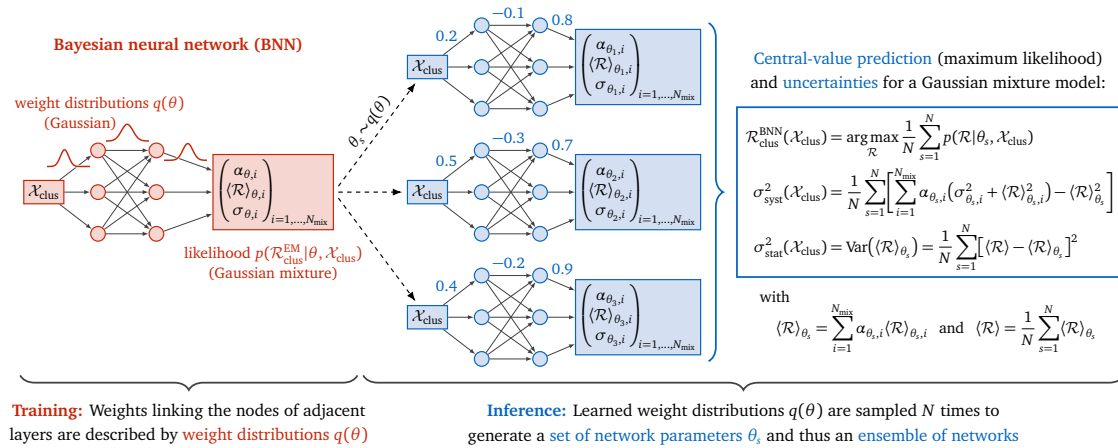


Figure 2: The principal design of the Bayesian neural network (BNN) employed for the regression fits and the corresponding uncertainty predictions for the topo-cluster calibration. The network on the left is designed such that the resulting weights linking the nodes of adjacent layers are described by weight functions $q(\theta)$ with learned averages and widths. The inference of the trained model, represented by a set of learned network parameters θ , is shown on the right, where an ensemble of N networks is sampled from $q(\theta)$ to generate N predictions from the model. The central prediction $\mathcal{R}_{\text{clus}}^{\text{BNN}}$ is the mode (most probable value) of the average of N probability density functions $p(\mathcal{R}|\theta_s, \mathcal{X}_{\text{clus}})$, where each individual $p(\mathcal{R}|\theta_s, \mathcal{X}_{\text{clus}})$ represents the probability that the network describes the target $\mathcal{R} = \mathcal{R}_{\text{clus}}^{\text{EM}}$ for a given topo-cluster with feature set $\mathcal{X}_{\text{clus}}$. Each of the $p(\mathcal{R}|\theta_s, \mathcal{X}_{\text{clus}})$ is defined as the mixture of $i = 1 \dots N_{\text{mix}}$ Gaussian distributions with learned means $\langle \mathcal{R} \rangle_{\theta_s, i}$, widths $\sigma_{\theta_s, i}$ and coefficients $\alpha_{\theta_s, i}$, as introduced in Eq. 12, with $N_{\text{mix}} = 3$. The averaged function is the normalised sum of all probability density functions obtained by sampling weights from $q(\theta)$ N times, with $N = 50$. It is composed of a total of $N \times N_{\text{mix}}$ Gaussian distributions. The weighted average response $\langle \mathcal{R} \rangle_{\theta_s}$ is the sum of $\alpha_{\theta_s, i} \langle \mathcal{R} \rangle_{\theta_s, i}$ from the N_{mix} distributions for each sample s . It is needed to predict the systematic ($\sigma_{\text{syst}}^{\text{BNN}}(\mathcal{X}_{\text{clus}})$) and statistical ($\sigma_{\text{stat}}^{\text{BNN}}(\mathcal{X}_{\text{clus}})$) uncertainties, which are further discussed in Section 5.2. It is calculated for each of the N sampled sets of network parameters θ_s to obtain $\langle \mathcal{R} \rangle_{\theta_s}$, the mean response averaged over the N corresponding $\langle \mathcal{R} \rangle_{\theta_s}$. This mean is needed exclusively for the calculation of $\sigma_{\text{stat}}^{\text{BNN}}$. All calculations are individually performed for any given topo-cluster at inference of the trained BNN model. The numbers written above some of the edges (links) between nodes are weights sampled from $q(\theta)$ at network inference. They are for illustration only and show the sampling character of the predictions.

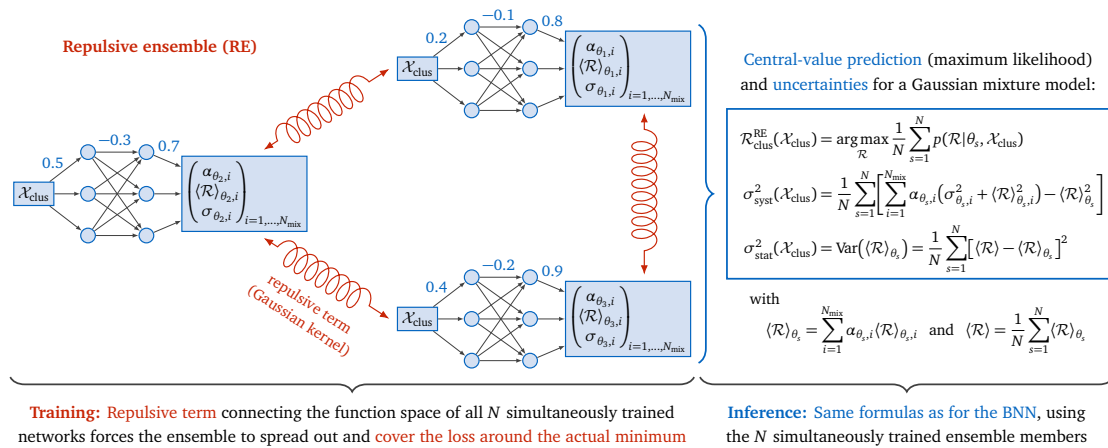


Figure 3: Schematics describing the design of the repulsive ensemble (RE) model employed for the regression fits and the corresponding uncertainty predictions of the topo-cluster calibration. An ensemble of N networks with identical configurations is configured. These N networks are trained simultaneously with interconnections between their loss functions acting as repulsive forces between them (shown as spring connectors in the schematic). The predictions from the N networks are collected for each topo-cluster in each processed batch during training. A central prediction and the uncertainties are calculated in the same way as illustrated in Figure 2 for the BNN, with the same nomenclature introduced there. The numbers written above some of the edges (links) between nodes illustrate the learned weights for one of the N repulsive ensembles. They are shown to illustrate that each network fits its own weights within the repulsive action of the loss function.

3.3.4 Repulsive ensemble loss function

The loss function of the RE network is derived from the modified update rule. Its principal form is a sum of $i = 1, \dots, N$ terms $\mathcal{L}_{\text{RE},i}$ for an ensemble of N networks,

$$\mathcal{L}_{\text{RE}} = \sum_{i=1}^N \mathcal{L}_{\text{RE},i}. \quad (14)$$

Each $\mathcal{L}_{\text{RE},i}$ has three components comprising

1. the negative logarithm of a posterior likelihood $p(D_{\text{train}}|\theta)$ of Bayes' theorem representing $p(\theta|D_{\text{train}})$, the likelihood of the network parameters θ to describe the training dataset D_{train} in Eq. 11, similar to the approach in the design of the BNN loss function in Section 3.3.2;
2. a repulsive term connecting the loss function space of all N networks with a given network $i \in [1, N]$ using Gaussian kernels;
3. and a Gaussian prior introduced by the application of Bayes' theorem when representing $p(\theta|D_{\text{train}})$ by $p(D_{\text{train}}|\theta)$, again similar to what is introduced for the BNN in Section 3.3.2.

The gradient of \mathcal{L}_{RE} is computed relative to all network parameters θ in the ensemble. The repulsive term in item 2 includes a *stop-gradient operation* on all θ in the ensemble, to assure valid gradients at each update. This term and the prior are also normalised by $1/N_{\text{train}}$ to avoid any scaling with the size of the training sample N_{train} . More insight on the formalism of the RE loss function is given in Appendix C.3.

4 Performance evaluations

The performance of the learned topo-cluster responses is fully evaluated for the BNN predictions ($\mathcal{R}_{\text{clus}}^{\text{BNN}}$) by applying the trained model to a test dataset that has no content overlap with the training dataset. Principal measures to determine the quality of this calibration are prediction power, the signal linearity and the local energy resolution. The median is used as a measure of the central tendency of the analysed distributions, because, unlike the mode, it is analytically accessible without explicitly determining (fitting) the shape of the distribution around the most probable value. The median is also relatively stable with respect to sample statistics, which can affect the statistical mean when distributions are not symmetric. Emerging tails in distributions of the prediction power and signal linearity, both introduced as evaluation metrics in Section 4.1 below, with increasing sample statistics were observed in some cases when evaluating them in bins of certain features or truth-related variables. These tails lead to observable changes of the means with increasing sample sizes in these cases, while medians were found to be stable once statistically meaningful samples were collected.

In addition, the performance of the BNN-derived topo-cluster calibration is compared to the one learned by the RE. Both calibrations are expected to produce the same central predictions and thus yield the same median signal linearity and energy resolution. This is a necessary requirement for a meaningful comparison of the uncertainties predicted by these two models.

4.1 Evaluation metrics

Determining the performance of the ML calibration cluster-by-cluster includes the evaluation of its accuracy or *prediction power*, similar to a closure test of a given functional model to predict the target in any calibration attempt, including non-machine-learning-based ones.⁹ In addition, the *signal linearity* achieved for the ML-calibrated energies are compared with the corresponding ones at other (calibrated) energy scales by determining their respective proportionality to the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the topo-cluster. Both prediction power and signal linearity are trivially connected for each topo-cluster, but show different distribution shapes for the test variables defined in Sections 4.1.1 and 4.1.2 below, leading to different measures of both centrality and width.

The energy resolution for individual topo-clusters is determined before any local calibration and after application of the ML- and the LCW-based calibrations. These corresponding comparisons allow identifying the best approach to mitigate calorimeter- and collision-environment-induced signal fluctuations affecting the cluster signal quality. The immediate expectation here is that calibration functions that are highly sensitive to the signal source yield the best resolution power.

The learned uncertainties are only available from the BNN and the RE network, and are further discussed in Section 5.

4.1.1 Accuracy of the learned response

To evaluate the accuracy of the ML calibrations, the test dataset is used to compare the outputs $\mathcal{R}_{\text{clus}}^{\kappa}$ of the networks κ for individual topo-clusters with the respective training target $\mathcal{R}_{\text{clus}}^{\text{EM}}$. The deviation from an accurate prediction $\Delta_{\mathcal{R}}^{\kappa}$ is introduced as a measure of the prediction power of the trained networks, with

$$\Delta_{\mathcal{R}}^{\kappa} = \frac{\mathcal{R}_{\text{clus}}^{\kappa}}{\mathcal{R}_{\text{clus}}^{\text{EM}}} - 1, \quad \text{with } \kappa \in \{\text{DNN}, \text{BNN}\}. \quad (15)$$

⁹Closure tests are usually performed using the same dataset with which the calibration was derived. In the context of the machine-learning-derived calibration, an independent dataset is used for this test.

It is calculated for each cluster in the test dataset with the expectation to approach zero for accurate predictions (highest prediction power). The prediction power of the DNN is taken from Ref. [4], which uses the same datasets.

4.1.2 Signal linearity

A fully calibrated calorimeter signal is expected to be directly proportional to a reference energy, which can be the incoming particle energy or, as in the case of the topo-cluster calibration, the locally deposited energy $E_{\text{clus}}^{\text{dep}}$ introduced in Section 2.3.2. The signal linearity can be measured cluster-by-cluster by Δ_E^κ , which quantifies its deviation from perfect linearity by¹⁰

$$\Delta_E^\kappa = \frac{E_{\text{clus}}^\kappa}{E_{\text{clus}}^{\text{dep}}} - 1, \quad \text{with } \kappa \in \{\text{EM, had, DNN, BNN}\}. \quad (16)$$

The considered topo-cluster energy scales κ are the basic EM scale given in Eq. 1, the hadronic (had) energy scale provided in the LCW sequence presented in Section 2.3.1, and the ML calibrations from the DNN [4] and the BNN, with the respective response predictions applied according to Eq. 6 in Section 3.2. The cluster signal at EM scale $E_{\text{clus}}^{\text{EM}}$ is not expected to be linear, due to the non-compensating character of the ATLAS calorimeters and the absence of corrections for that at this scale. Larger deviations of Δ_E^{EM} from zero are expected for small $E_{\text{clus}}^{\text{dep}}$, where ionisations increasingly dominate the energy loss, and in regimes where pile-up significantly adds to $E_{\text{clus}}^{\text{EM}}$. Therefore, $E_{\text{clus}}^{\text{EM}}$ is not considered to be calibrated, unlike all other considered energy scales, in the context of the results presented here. Nevertheless, it provides a baseline for the signal-related topo-cluster performance.

Calibrated energy scales are supposed to provide signals directly proportional to $E_{\text{clus}}^{\text{dep}}$. A residual level of fluctuations can be expected after topo-cluster calibration, due to the highly stochastic nature of its response. Nevertheless, the chosen measure of centrality, the median deviation from linearity $\langle \Delta_E \rangle_{\text{med}}$, is expected to be zero after calibrations are applied. Any deviation from this expectation can affect the measurement of final state objects like particles, jets and event shapes composed of topo-clusters in terms of the introduction of kinematic biases, an increase of systematic uncertainties and a potential degradation of the resolution.

4.1.3 Relative local energy resolution

The measure of the relative local energy resolution $\sigma_{\text{rel}}^E(E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}})$ employs the 68% inter-quantile range $Q_{f=68\%}^w$ and the median of the respective $\Delta_E^\kappa + 1 = E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}}$ distribution,

$$\sigma_{\text{rel}}^E(E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}}) = \frac{Q_{f=68\%}^w(E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}})}{2\langle E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}} \rangle_{\text{med}}}, \quad \text{with } \kappa \in \{\text{EM, had, DNN, BNN}\}, \quad (17)$$

on the various energy scales κ . Its construction is based on Eqs. D.1 and D.3 given in Appendix D.1.

¹⁰The prediction power represented by $\Delta_{\mathcal{R}}$ and the signal linearity represented by Δ_E are related by the response definition given in Eq. 4 and the reconstruction of the calibrated energy given in Eq. 6, such that $\Delta_E = -\Delta_{\mathcal{R}}/(\Delta_{\mathcal{R}} + 1)$. This generally leads to two different distributions with possibly different central values, $\langle \Delta_{\mathcal{R}} \rangle_{\text{med}} \neq \langle \Delta_E \rangle_{\text{med}}$, and widths for any selected sample of topo-clusters, and thus to differences in the performance. Only when the response is perfectly learned, indicated by $\Delta_{\mathcal{R}} = 0$, a perfect signal linearity $\Delta_E = 0$ is achieved for a given topo-cluster.

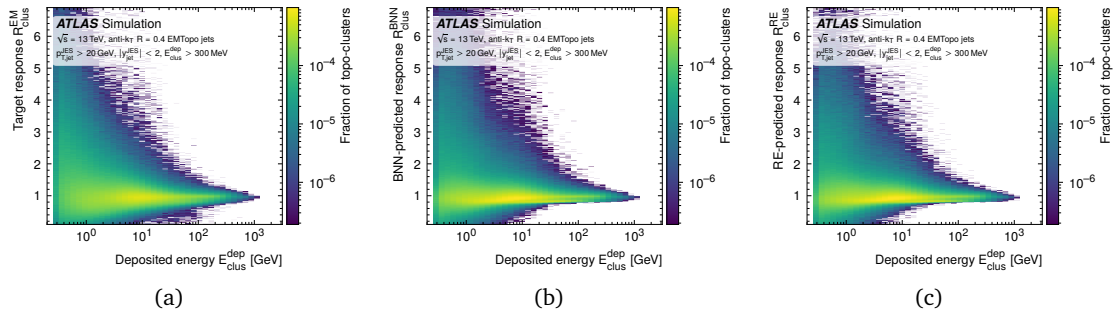


Figure 4: The topo-cluster response $\mathcal{R}_{\text{clus}}$ is evaluated cluster-by-cluster as a function of the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the cluster. In (a), the distribution of the response at EM scale $\mathcal{R}_{\text{clus}}^{\text{EM}}$ (the training target) is shown, while (b) shows the distribution of the response $\mathcal{R}_{\text{clus}}^{\text{BNN}}$ predicted by the trained BNN. The corresponding distribution of the response $\mathcal{R}_{\text{clus}}^{\text{RE}}$ predicted by the RE is shown in (c). The topo-clusters are extracted from MC simulations, as described in Section 3.1.1.

4.2 Prediction power

The prediction power of the ML-learned topo-cluster calibration can be evaluated cluster-by-cluster as functions of various variables available from MC simulations and as functions of any of the features from the feature set $\mathcal{X}_{\text{clus}}$ used for the training, for both the BNN-derived and the RE-derived calibrations. A qualitative analysis allows comparing distribution shapes for the predictions from the networks with the corresponding distributions for the target, without the need to introduce statistical measures that may not describe the overall distribution well. The quantitative analysis introduces measures for the central values of these distributions that allow comparing them directly using, e.g., $\Delta_{\mathcal{R}}$ from Eq. 15.

4.2.1 Distribution shapes

A qualitative comparison of the distributions of the training target $\mathcal{R}_{\text{clus}}^{\text{EM}}$ and the predictions from both the BNN and the RE, all as a function of the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the topo-cluster, is shown in Figure 4. The target distribution in Figure 4(a) is reproduced well by both the BNN (Figure 4(b)) and RE (Figure 4(c)) predictions. Some small shifts in the populations between target and predictions can be observed in that the networks provide a sharper image of regions of higher population density. These indicate that the learned $\mathcal{R}_{\text{clus}}^{\text{EM}}$ is dominated by the statistically better illuminated regions, thus giving a qualitative impression of the limitations of the networks to predict outliers. With this respect, both the BNN and the RE network show very similar behaviour.

Further qualitative analyses of the BNN and RE predictions as functions of selected features used for the network training of the topo-cluster response are collected in Figure 5. As already observed in Figures 4(b) and 4(c), the learned responses show some differences at the edges of the densely populated ridges when compared with the corresponding target distributions and among themselves, and evaluated as a function of $E_{\text{clus}}^{\text{EM}}$. This again is a display of *statistical dominance* of the distribution regions with higher population densities in the derivation of the predictions.

The complex distributions of $\mathcal{R}_{\text{clus}}^{\text{EM}}$ as a function of the cluster signal fraction in the electromagnetic calorimeter f_{emc} shown in Figure 5(d) arise from the response characteristics and the calorimeter geometry. The near horizontal ridge extending from $\mathcal{R}_{\text{clus}}^{\text{EM}} \approx 0.8$ at $f_{\text{emc}} = 0$ to $\mathcal{R}_{\text{clus}}^{\text{EM}} \approx 1.0$ for $f_{\text{emc}} = 1$ shows the expected behaviour for a non-compensating

calorimeter where a lower f_{emc} is indicative to an increased contribution from hadronic showers in topo-clusters exclusively located in the hadronic calorimeter or extending from the electromagnetic to the hadronic calorimeter. The upper ridge is populated by topo-clusters with rising $\mathcal{R}_{\text{clus}}^{\text{EM}}$ for decreasing f_{emc} . It is indicative of topo-clusters with $E_{\text{clus}}^{\text{dep}}$ predominantly generated by ionisations in hadronic showers that increasingly extend into the Tile calorimeters. The cluster signal $E_{\text{clus}}^{\text{EM}}$ at EM scale overestimates $E_{\text{clus}}^{\text{dep}}$ because these ionizations generate larger signals than electromagnetic showers depositing the same $E_{\text{clus}}^{\text{dep}}$ in this calorimeter [40, 41].

Figures 5(g)–(i) show the distributions of $\mathcal{R}_{\text{clus}}^{\text{EM}}$, $\mathcal{R}_{\text{clus}}^{\text{BNN}}$ and $\mathcal{R}_{\text{clus}}^{\text{RE}}$ as a function of the distance of the topo-cluster centre-of-gravity $|\vec{c}_{\text{clus}}|$ from the nominal vertex in ATLAS. The most prominent observed structure around $|\vec{c}_{\text{clus}}| \approx 3900$ mm is introduced by the detector geometry. It shows the effect of the transition from the central barrel to the endcap calorimeters.

The general observation from the comparison of the $\mathcal{R}_{\text{clus}}^{\text{EM}}(f_{\text{emc}})$ distribution in Figures 5(d)–(f) and the $\mathcal{R}_{\text{clus}}^{\text{EM}}(|\vec{c}_{\text{clus}}|)$ distribution in Figures 5(g)–(i) is that the complex structures of the target distributions are well reproduced by both the BNN and the RE predictions, with some slightly sharpened edges around the ridges of higher topo-cluster populations for the predictions due to the statistical dominance of these regions in the learning.

4.2.2 Central prediction power

Results from a quantitative analysis of the prediction power of the BNN topo-cluster calibration model are collected in Figure 6. The metric used here is the relative difference $\Delta_{\mathcal{R}}$ between learned $\mathcal{R}_{\text{clus}}^{\text{BNN}}$ and the target $\mathcal{R}_{\text{clus}}^{\text{EM}}$, as defined in Eq. 15 in Section 4.1.1. It is measured cluster-by-cluster and analysed as functions of variables from MC simulations and the features in $\mathcal{X}_{\text{clus}}$, using two different measures of centrality for the resulting $\Delta_{\mathcal{R}}$ distributions. For example, Figure 6 shows the mean and median prediction power of the BNN as function of $E_{\text{clus}}^{\text{dep}}$, $E_{\text{clus}}^{\text{EM}}$, f_{emc} and $|\vec{c}_{\text{clus}}|$ and compares those to the median and mean prediction power of the DNN [4]. The expectation from the particular (different) choices of loss functions for the BNN and the DNN training is that the mode $\langle \Delta_{\mathcal{R}} \rangle_{\text{mode}} \approx 0$ for all $\Delta_{\mathcal{R}}$ distributions in bins of features. It is found that $\langle \Delta_{\mathcal{R}} \rangle_{\text{med}} \approx \langle \Delta_{\mathcal{R}} \rangle_{\text{mode}}$ for both networks and all evaluation scales, with the advantage that $\langle \Delta_{\mathcal{R}} \rangle_{\text{med}}$ is well defined and does not require a particular algorithm or peak shape fit to be determined. The findings presented in Figure 6 show that the BNN has a similar if not stronger median prediction power than the DNN for all shown evaluation scales with the exception of $E_{\text{clus}}^{\text{EM}}$. The mean prediction power is, as expected from the choice of the loss functions for the BNN and the DNN, not as strong as the median for all evaluation scales.

4.2.3 Network training issues

About 3% of all topo-clusters collected from the simulated calorimeter jets were found to be resistant to the BNN, RE, and DNN training. These clusters are characterised by a target response $\mathcal{R}_{\text{clus}}^{\text{EM}} \gg 1$. This resistance potentially arises from a large pile-up contribution that is not accounted for in $E_{\text{clus}}^{\text{dep}}$ but increases $E_{\text{clus}}^{\text{EM}}$ significantly. As the affected topo-clusters are mostly observed in the hadronic calorimeters with considerably coarser readout geometry than the electromagnetic ones, this signal contribution from pile-up at the cluster location may not be reflected in the features due to a limited (spatial) resolution. The consequential lack of precision of the features can lead to a loss of expressiveness concerning the particular topo-cluster response and thus a severely reduced prediction power. In addition to the location, this interpretation is supported by the frequency of occurrence of these training-resistant topo-clusters as a function of pile-up, rising from about 0.2% of all clusters collected at $\mu \simeq 0$ (no pile-up) to close to 10% of all clusters collected at $\mu \simeq 90$ (highest observed pile-up level in the dataset).

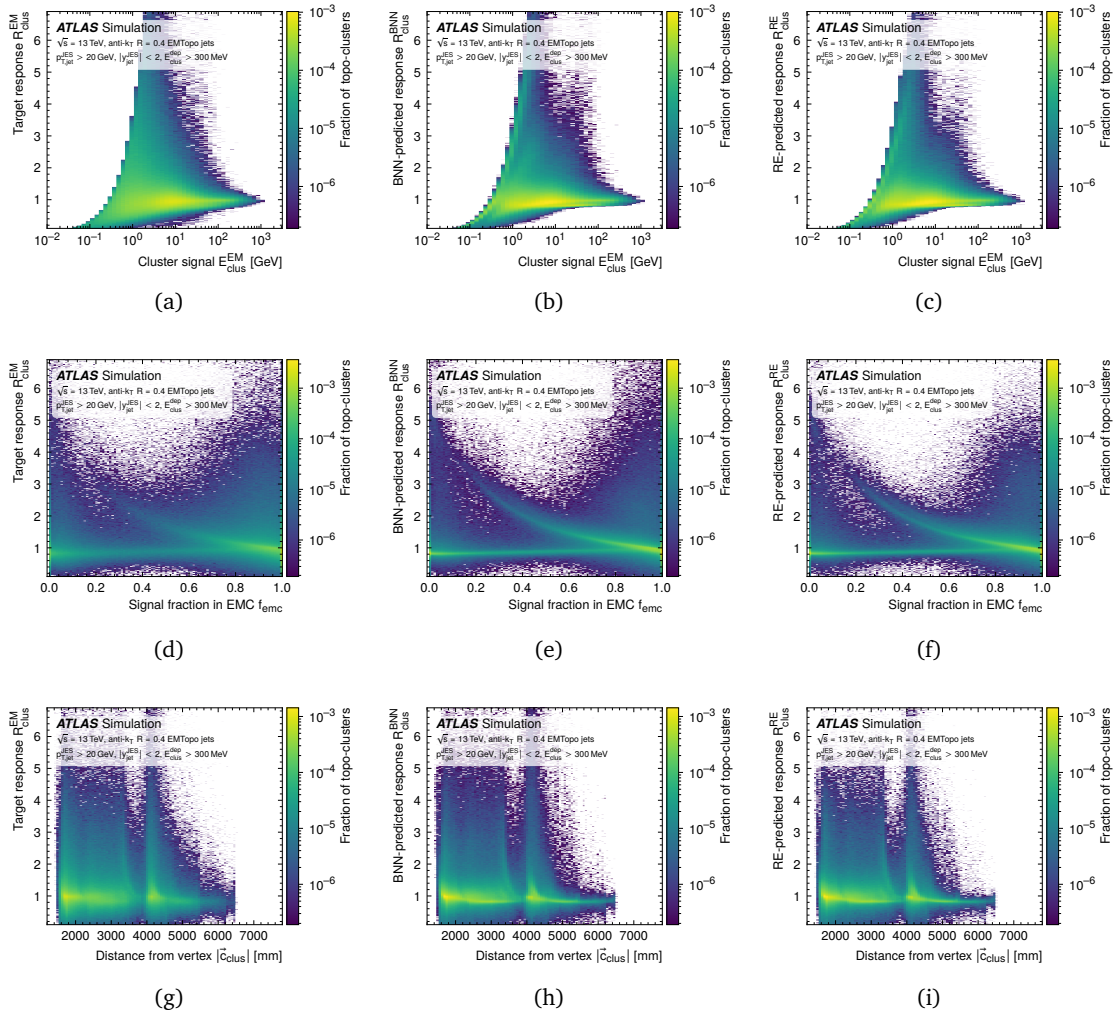


Figure 5: The distributions of the topo-cluster response $\mathcal{R}_{\text{clus}}^{\text{EM}}$ and the corresponding predictions from the BNN ($\mathcal{R}_{\text{clus}}^{\text{BNN}}$) and the RE ($\mathcal{R}_{\text{clus}}^{\text{RE}}$) are evaluated cluster-by-cluster as a function of selected features. The distributions of the training target $\mathcal{R}_{\text{clus}}^{\text{EM}}$ are shown in (a) as a function of the cluster signal $E_{\text{clus}}^{\text{EM}}$, in (d) as a function of the cluster signal fraction f_{emc} in the electromagnetic calorimeter, and in (g) as a function of the distance of the topo-cluster centre-of-gravity $|\vec{c}_{\text{clus}}|$. The distributions of the predicted responses $\mathcal{R}_{\text{clus}}^{\text{BNN}}$ and $\mathcal{R}_{\text{clus}}^{\text{RE}}$ as functions of the same features are respectively shown in (b), (c) for $E_{\text{clus}}^{\text{EM}}$, in (e), (f) for f_{emc} , and in (h), (i) for $|\vec{c}_{\text{clus}}|$. The topo-clusters are extracted from MC simulations, as described in Section 3.1.1.

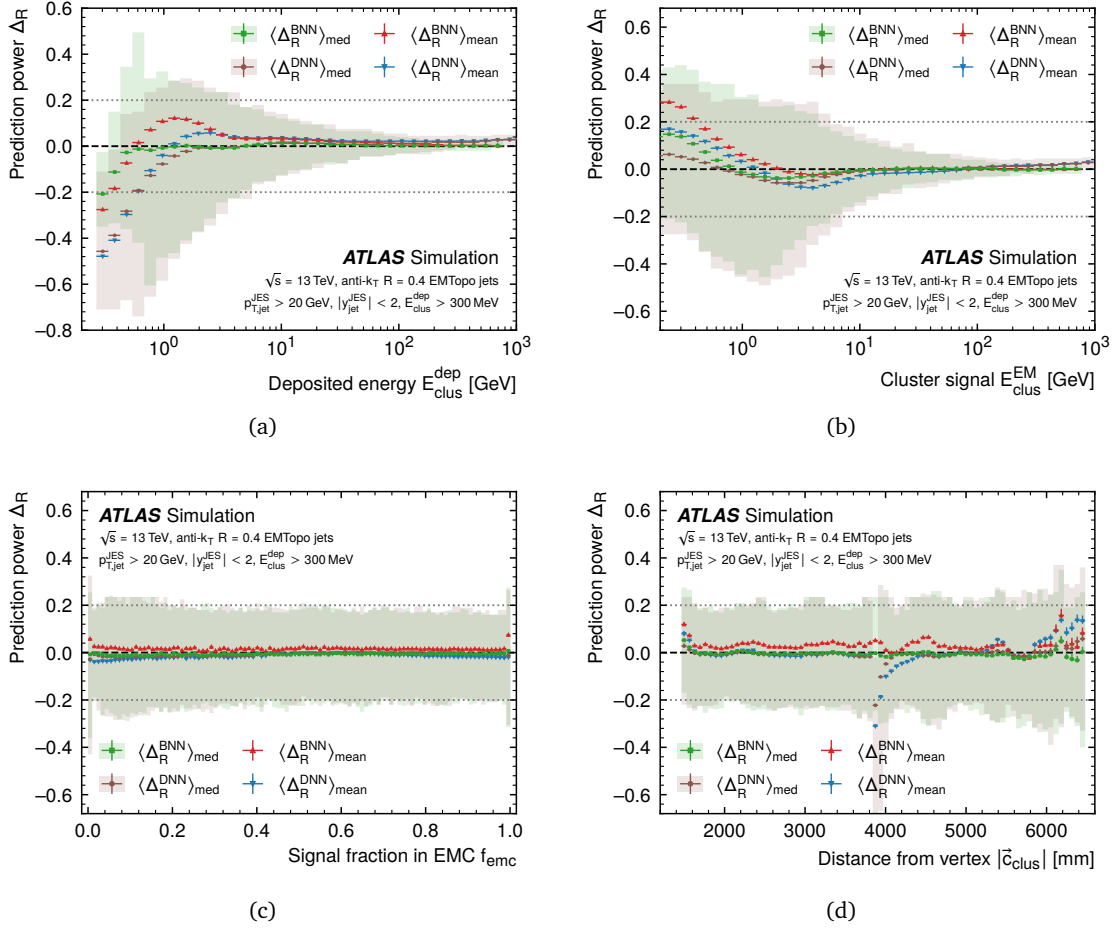


Figure 6: The mean and median prediction power of the trained BNN ($\Delta_{\mathcal{R}}^{\text{BNN}}$) and the trained DNN ($\Delta_{\mathcal{R}}^{\text{DNN}}$) are compared as a function of (a) the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the topo-cluster, (b) the cluster signal $E_{\text{clus}}^{\text{EM}}$, (c) the cluster signal fraction in the electromagnetic calorimeter f_{emc} , and (d) the cluster centre-of-gravity $|\vec{c}_{\text{clus}}|$. The measures of centrality for the BNN are determined from the corresponding distributions shown in Figure 4(b) and in Figure 5(b), (e) and (h), while the ones for the DNN are extracted from Ref. [4]. The topo-clusters are collected from MC simulations, as described in Section 3.1.1. The shaded areas indicate the narrowest range of the respective $\Delta_{\mathcal{R}}$ distributions containing 68% of all entries. Statistical uncertainties are indicated by vertical bars if larger than the symbol size.

A dedicated neural network was trained to tag the affected topo-clusters and remove them from the training and test datasets for the results presented in this paper. The ability to efficiently tag them, and to locate them in certain calorimeter regions under certain pile-up conditions, supports the potential application of dedicated calibrations and corrections beyond the scope of the study presented here. Further explorations are under way to address these issues.

4.3 Central calibration performance

The prediction power discussed above is not a sufficient indicator for the accuracy and precision of the corresponding topo-cluster calibration. The deviation from signal linearity measure Δ_E introduced in Section 4.1.2 provides a direct scale to determine both, and thus the overall quality of the learned calibration. This measure can be constructed at various energy scales. For comparisons, the median deviation from linearity is evaluated for topo-clusters from the test dataset as a function of $E_{\text{clus}}^{\text{dep}}$ and selected features for the basic EM scale (Δ_E^{EM}), the LCW hadronic scale (Δ_E^{had}), the energy scale derived from the DNN calibration (Δ_E^{DNN}), and the scale after applying the BNN-learned calibration (Δ_E^{BNN}).

Figure 7 compares Δ_E on the various energy scales, as a function of selected topo-cluster features and $E_{\text{clus}}^{\text{dep}}$. As seen in Figure 7(a), the BNN-derived calibration extends the region of accurate energy reconstruction with $|\langle \Delta_E^{\text{BNN}} \rangle_{\text{med}}| \lesssim 2\%$ for $E_{\text{clus}}^{\text{dep}} \gtrsim 500$ MeV down from the next-best calibration based on the DNN predictions, where a similar median deviation from linearity is reached for $E_{\text{clus}}^{\text{dep}} \gtrsim 1$ GeV. This is a significant gain of accuracy in the low energy regime.

When evaluated as a function of the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$, the accuracy of both ML calibrations is significantly improved compared to the standard LCW-based hadronic calibration and the EM scale energy reconstruction, as is shown in Figure 7(b). The accuracy is not as high as observed in Figure 7(a), which is likely due to a residual, unmitigated and highly stochastic contribution to smaller signals that is mostly introduced by pile-up dominating $E_{\text{clus}}^{\text{EM}}$ in this region. Nevertheless, both machine-learning based calibrations show significant less deviation from linearity than the standard LCW hadronic energy scale, which is derived from a calibration that uses $E_{\text{clus}}^{\text{EM}}$ as input as well.

Both $E_{\text{clus}}^{\text{BNN}}$ and $E_{\text{clus}}^{\text{DNN}}$ provide an accurate median measure of $E_{\text{clus}}^{\text{dep}}$ across the illuminated cluster rapidity $y_{\text{clus}}^{\text{EM}}$ space, as shown in Figure 7(c). Some increase of both $\langle \Delta_E^{\text{BNN}} \rangle_{\text{med}}$ and $\langle \Delta_E^{\text{DNN}} \rangle_{\text{med}}$ can be observed at the boundaries of this space at $|y_{\text{clus}}^{\text{EM}}| \approx 2.5$. The topo-clusters reconstructed at these edges can extend beyond the $|\eta_{\text{det}}| = 2.5$ boundary, where the cell readout granularity of the ATLAS calorimeters is reduced by a factor of four. The increase in cell size potentially reduces the expressiveness of some features for these clusters. This leads to a loss of central accuracy, mainly due to a loss of sensitivity of the affected features that mitigate stochastic effects on $\mathcal{R}_{\text{clus}}^{\text{EM}}$. A larger deviation from linearity measured by $\langle \Delta_E^{\text{had}} \rangle_{\text{med}}$ at the hadronic LCW scale, already observed in Figure 7(b) for $\langle \Delta_E^{\text{had}} \rangle_{\text{med}}(E_{\text{clus}}^{\text{EM}})$, is seen for $\langle \Delta_E^{\text{had}} \rangle_{\text{med}}(y_{\text{clus}}^{\text{EM}})$ in Figure 7(c) as well for $|y_{\text{clus}}^{\text{EM}}| \lesssim 2$, with maxima for topo-clusters located in the complex calorimeter transition region at $|\eta_{\text{det}}| \approx 1.5$. Both BNN and DNN learned calibrations measure $E_{\text{clus}}^{\text{dep}}$ in this transition region well, indicating that the feature expressiveness is sufficient for an accurate calibration.

Figure 7(d) shows the significant deviation from linearity for the basic topo-cluster signal $E_{\text{clus}}^{\text{EM}}$, and similarly for $E_{\text{clus}}^{\text{had}}$, for topo-cluster signal times t_{clus} reconstructed outside of the bunch crossing time window of $-12.5 \text{ ns} < t_{\text{clus}} < 12.5 \text{ ns}$. In particular $\langle \Delta_E^{\text{EM}} \rangle_{\text{med}}(t_{\text{clus}}) > 1$ hints on the cluster response $\mathcal{R}_{\text{clus}}^{\text{EM}}$ being dominated by out-of-time pile-up signal contributions. This leads to a slightly increased $\langle \Delta_E^{\text{had}} \rangle_{\text{med}}(t_{\text{clus}}) > \langle \Delta_E^{\text{EM}} \rangle_{\text{med}}(t_{\text{clus}})$ for $|t_{\text{clus}}| \gtrsim 12.5 \text{ ns}$, as the hadronic calibration from LCW is agnostic to t_{clus} .

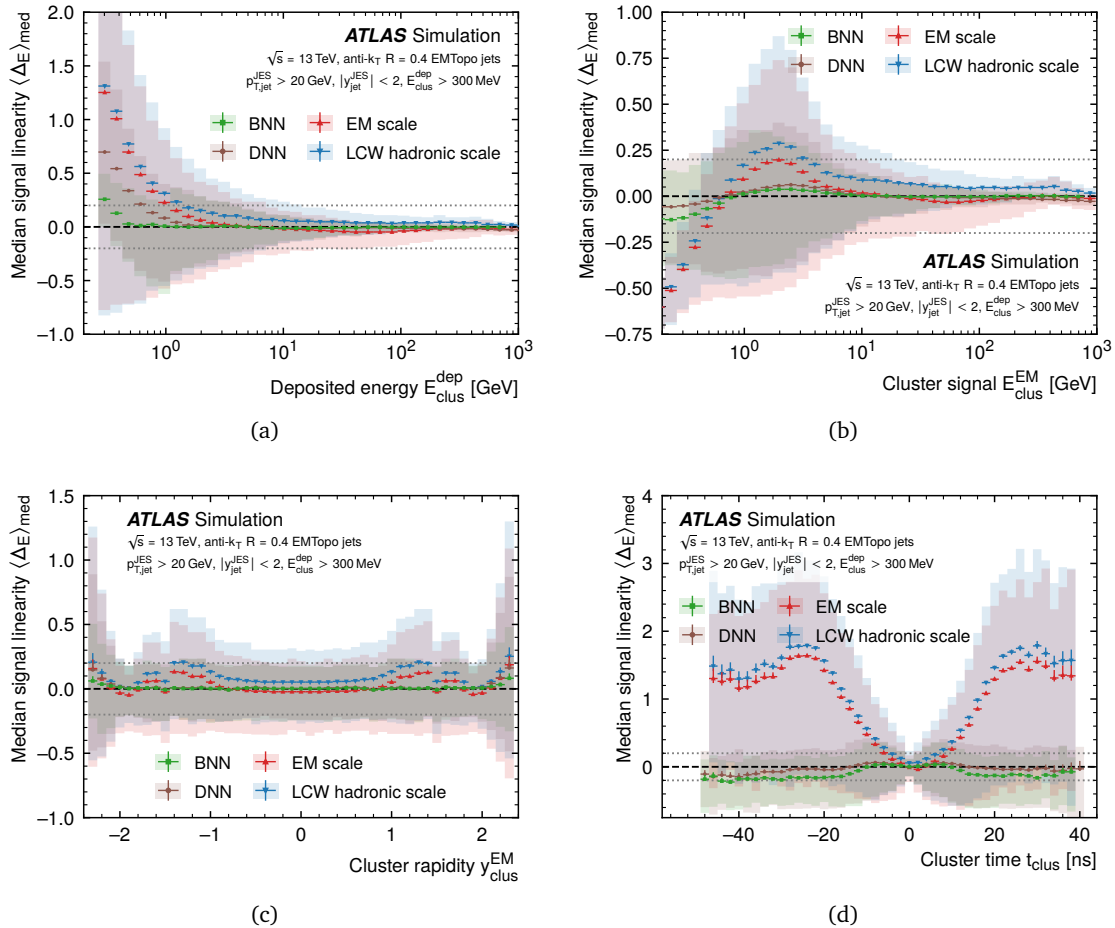


Figure 7: The median deviation from linearity $\langle \Delta_E \rangle_{\text{med}}$ for topo-clusters at the basic EM, the LCW hadronic, the BNN-derived and the DNN-derived energy scales is shown as a function of (a) the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the cluster, (b) the cluster signal $E_{\text{clus}}^{\text{EM}}$, (c) the cluster rapidity reconstructed at the basic EM scale $y_{\text{clus}}^{\text{EM}}$, and (d) the cluster signal time t_{clus} . The topo-clusters are extracted from MC simulations according to the description in Section 3.1.1. The shaded areas indicate the narrowest range of the respective Δ_E distributions containing 68% of all entries. Statistical uncertainties are indicated by vertical bars if larger than the symbol size.

Both BNN and DNN trained calibrations learn $\mathcal{R}_{\text{clus}}^{\text{EM}}(t_{\text{clus}})$ well, which leads to $\langle \Delta_E^{\text{BNN}} \rangle_{\text{med}}(|t_{\text{clus}}| < 12.5 \text{ ns}) \approx 0$ and $\langle \Delta_E^{\text{DNN}} \rangle_{\text{med}}(|t_{\text{clus}}| < 12.5 \text{ ns}) \approx 0$. Outside of this time window, the BNN-derived calibration shows a near-constant median deviation from linearity by underestimating $E_{\text{clus}}^{\text{dep}}$ by about 20%. The DNN-derived calibration shows a better performance for $t_{\text{clus}} > 12.5 \text{ ns}$, with a deviation from linearity within a few percent, indicating that the signal contributions from the following bunch crossings are mitigated well.

Most of the features in $\mathcal{X}_{\text{clus}}$ are not involved in the reconstruction of both $E_{\text{clus}}^{\text{EM}}$ and $E_{\text{clus}}^{\text{had}}$, at least not directly as inputs. Extending the dimensionality of the local topo-cluster calibration, which is well supported by the machine learning applications, allows inclusion of information reflecting the cluster source and the pp collision environment. In addition to the results presented in Figure 7, the deviation from linearity is evaluated cluster-by-cluster for each feature in the full feature space represented in the test dataset. Figure 8 shows a selection of results for some of the features that both the EM scale and the hadronic scale from LCW are agnostic to. The results for the remaining features are collected in Appendix E.1.

A median calibration accuracy of better than 2% is achieved with both the BNN and the DNN nearly everywhere, when evaluated over the whole f_{emc} , $|\vec{c}_{\text{clus}}|$, N_{pV} and μ value ranges. This indicates that both networks learn the response well from these features, even if those features are distributed in a complex way, as seen for f_{emc} and $|\vec{c}_{\text{clus}}|$ in Figure 6(c) and 6(d), for example. In addition, the presence of a significant pile-up activity that increases the stochasticity of the topo-cluster signal and its features is well mitigated when the corresponding pile-up measures N_{pV} and μ are included in the (ML-learned) calibration.

4.4 Local topo-cluster energy resolution

The relative local energy resolution of topo-clusters is evaluated at the basic EM, the hadronic scale from LCW, and for the BNN- and DNN-derived calibrations. It is measured using $\sigma_{\text{rel}}^E(E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}})$ introduced in Eq. 17 in Section 4.1.3, where κ names the energy scale.

Figure 9 shows $\sigma_{\text{rel}}^E(E_{\text{clus}}^\kappa/E_{\text{clus}}^{\text{dep}})$ as a function of $E_{\text{clus}}^{\text{dep}}$, $E_{\text{clus}}^{\text{EM}}$, and the pile-up measures N_{pV} and μ . Both ML calibrations yield a significantly better relative resolution than the standard local hadronic calibration, together with a significant improvement relative to the EM scale resolution, on all evaluation scales. This is expected from the larger number of features entering the BNN and the DNN and the resulting increased exposure of these calibrations to more information concerning the topo-cluster signal source, in combination with the exploitation of feature correlations lost for other energy scale calibrations, and the learned smooth calibration functions.

The steeper downturn of the calibrated signal fluctuations for the BNN-derived calibration observed in Figure 9(a) for $E_{\text{clus}}^{\text{dep}} \lesssim 1 \text{ GeV}$ indicates that this model best learns the transition of the signal source from inelastic hadronic interactions to more ionisation-dominated signals.¹¹ This is likely associated with the use of a Gaussian mixture model for $p(D_{\text{train}}|\theta)$ in the BNN loss function, as described in Eq. 12 of Section 3.3.2, among other possible differences between the BNN and the DNN. The comparisons of the individual shapes of the Δ_E^{BNN} and the Δ_E^{DNN} distributions in bins of $E_{\text{clus}}^{\text{dep}}$ show significantly suppressed tails in case of the BNN-derived calibrations. This is expected for signals from ionisation losses, which are subject to considerably less intrinsic fluctuations than signals from energy loss mechanisms occurring in inelastic processes. A representative selection of these comparisons is shown in Figure 19 in Appendix E.2.

¹¹Hadrons can traverse relatively large distances in matter, if their energy is below the typical threshold of about 1 GeV for inelastic interactions, according to Ref. [42]. This suggests that energy deposits in topo-clusters within $0 < E_{\text{clus}}^{\text{dep}} \lesssim 1 \text{ GeV}$ reflect a significant contribution to $E_{\text{clus}}^{\text{dep}}$ from energy loss mechanisms not involving hadronic inelastic interactions, like ionisations by charged particles.

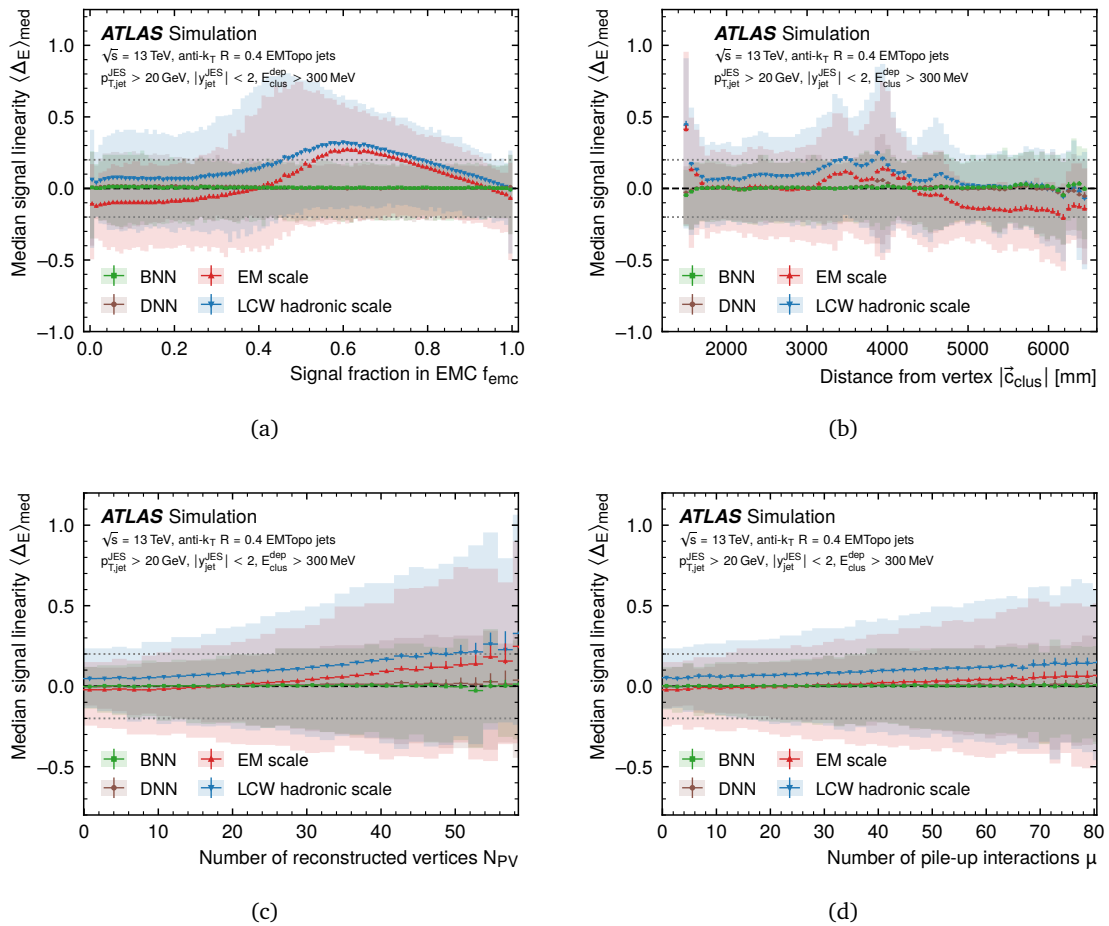


Figure 8: The median deviation from signal linearity $\langle \Delta_E \rangle_{\text{med}}$ for topo-cluster signals at the basic EM, the LCW hadronic, the BNN-derived and the DNN-derived energy scales is shown as a function of (a) the cluster signal fraction in the electromagnetic calorimeter f_{emc} , (b) the distance of the cluster centre-of-gravity from the nominal collision vertex $|\vec{c}_{\text{clus}}|$, (c) the number of reconstructed primary vertices N_{PV} (in-time pile-up measure), and (d) the number of pile-up collisions in the bunch crossing μ (out-of-time pile-up measure). The results are achieved with topo-clusters from the test dataset, extracted as described in Section 3.1.1. The shaded areas indicate the narrowest range of the respective Δ_E distributions containing 68% of all entries. Statistical uncertainties are indicated by vertical bars if larger than the symbol size.

The relative topo-cluster energy resolution after application of the ML-based calibrations is increased, compared to the corresponding resolutions obtained for the standard energy scales, when analysed as a function of the in-time pile-up measure N_{PV} in Figure 9(c), and as a function of the out-of-time pile-up measure μ in Figure 9(d). The increased but still partial cluster-by-cluster pile-up mitigation is manifested by the reduced relative signal fluctuations observed at any given N_{PV} and μ when compared with the other energy scales. The fluctuations after applying the BNN-derived topo-cluster calibration show a more flat dependence on N_{PV} than the residual fluctuations after application of the DNN-based calibration, with a slightly worse performance at low N_{PV} (low pile-up) but a better performance at higher N_{PV} (higher pile-up). A similar, but less pronounced behaviour is observed in Figure 9(d), where the relative energy resolutions achieved with the BNN- and DNN-derived calibrations are shown as a function of μ .

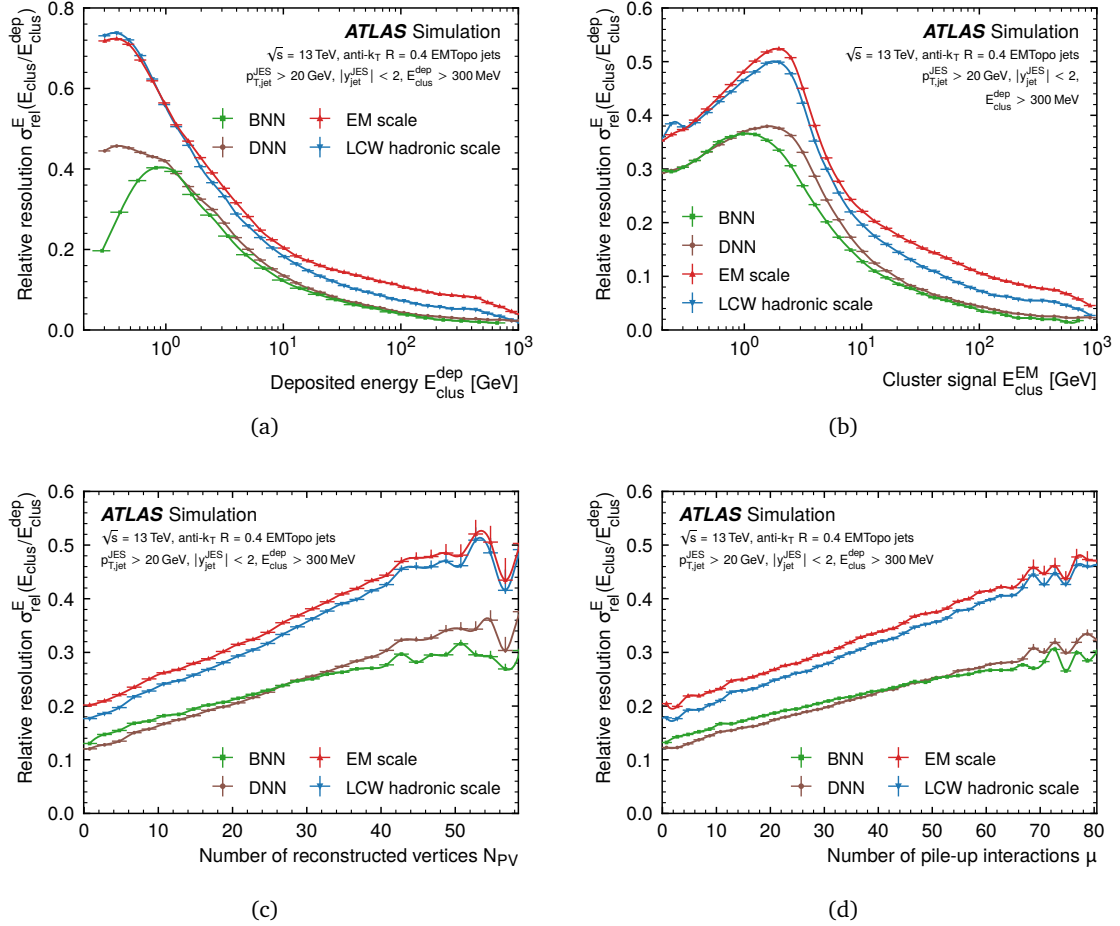


Figure 9: The relative local energy resolution of topo-clusters at the basic EM, the LCW hadronic, the BNN-derived and the DNN-derived energy scales, measured using Eq. 17, as a function of (a) the energy deposited in the cluster E_{clus}^{dep} , (b) the basic cluster signal E_{clus}^{EM} , (c) the number of reconstructed vertices N_{PV} , and (d) the number of pile-up collisions μ in the bunch crossing. The topo-clusters are extracted from the test dataset as discussed in Section 3.1.1. Statistical uncertainties are indicated by vertical bars if larger than the symbol size.

Table 4: The principal qualitative relation between aleatoric and epistemic predictive uncertainties in machine-learning applications and the systematic and statistical uncertainty predictions for the topo-cluster calibration. These relations are context-driven.

Predictive uncertainty	Uncertainty component	Mitigation strategy
<i>Epistemic (model) uncertainties</i>		
Network architecture	systematic	network configuration optimisation
Hyper-parameter choice	systematic	network configuration optimisation
Lack of expressivity of data	systematic	feature choices
Insufficient amount of data	statistical	sufficiently large training dataset
<i>Aleatoric (data) uncertainties</i>		
Residual level of stochasticity in data	systematic	none (irreducible noise)

5 Analyzing the uncertainty predictions

The uncertainty predictions from both the BNN and the RE are based on predictions sampled from a model representing an ensemble of trained networks. In case of the BNN, this ensemble is formed by sampling the network parameters at inference from variational distributions with learned parameters, as schematically laid out in Figure 2 in Section 3.3.1. For the RE, the ensemble is a collection of networks interconnected by a repulsive term during training, as discussed in Section 3.3.3.

5.1 Predictive uncertainties

Learned uncertainties are usually referred to as *predictive uncertainties* [10, 43, 44]. They are divided into two major categories,

1. *epistemic* or *model uncertainties* that are related to how machines can learn and what cannot be learned, and
2. *aleatoric* or *data uncertainties* that are related to the residual level of stochasticity in the data that cannot be mitigated any further by the model.

The expectation is that the epistemic uncertainties, after the final network optimisation for the topo-cluster calibration, are dominated by contributions inherent to the data that limit what the BNN can learn. In particular, contributions to these uncertainties are possible from a statistically insufficient coverage of the whole feature space, or certain regions of it, by topo-clusters in the training dataset. These contributions are thus expected to largely vanish in the limit of infinite training data and a sufficient number of epochs. As said above, the aleatoric uncertainties reflect the lack of expressiveness of the data introduced by a relatively high level of noise predominantly introduced by pile-up that cannot be overcome by increasing the training dataset or improving the network model.

The predictive uncertainty of a trained network model is a measure for its lack of expressive power arising from either of the two categories of uncertainties introduced above. This

ambiguity can be resolved by relating the learned uncertainties to the standard categories of statistical and systematic uncertainties that are usually associated with the application of a calibration model, or with a measurement in general. The chosen convention here follows the general expectation that the statistical uncertainty vanishes in the limit of infinite training data, and the systematic uncertainty remains finite at this limit. When evaluating the quality of the network, both the statistical and the systematic uncertainty contribute to the learned uncertainty as a systematic component in the sense that evaluating the network many times, without changing the training dataset size, does not reduce them. They are thus typically considered as external nuisance parameters associated with the specific network configuration and training conditions. The principal contributions of the predictive uncertainties to the uncertainty predictions associated with the topo-cluster calibration are collected in Table 4.

5.2 Uncertainty components

Separately analysing the statistical (σ_{stat}) and systematic (σ_{syst}) component of the total learned uncertainty σ_{tot} requires a sufficiently large training dataset. It is possible for both the BNN and the RE network employed here, as they use the same large training dataset introduced in Table 3 of Section 3.3. This means that

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{syst}}^2 + \sigma_{\text{stat}}^2}. \tag{18}$$

Following the previous discussion, σ_{syst} and σ_{stat} are characterised by their behaviour in the limit of an infinite training dataset size N_{train} ,

$$N_{\text{train}} \rightarrow \infty \Rightarrow \begin{cases} \sigma_{\text{syst}} \rightarrow \text{const.} > 0, \\ \sigma_{\text{stat}} \rightarrow 0. \end{cases}$$

With this understanding, σ_{syst} represents the limit of the accuracy to which the model learns to describe the target with the training data available. In case of a non-deterministic or stochastic target (noisy data) [12], like $\mathcal{R}_{\text{clus}}^{\text{EM}}$, limited expressiveness of the network (structure uncertainty) [13], and non-optimal choices of hyper-parameters (uncertainty in model parameters) can all be interpreted in the spirit of contributions to the systematic uncertainty inherent to the calibration model and the training data. It has contributions to both epistemic and aleatoric uncertainty components, as qualitatively summarised in Table 4.

The statistical uncertainty component σ_{stat} is constructed to become insignificantly small with increasing N_{train} , and is thus expected to become irrelevant for a sufficiently large training dataset. Therefore, σ_{stat} can be considered as a significant part of the epistemic uncertainties. A brief and more formal introduction to the calculation of σ_{syst} and σ_{stat} can be found in Appendix D.2.

The asymptotic behaviours of the uncertainty components as a function of the training dataset size are summarised in Figure 10. The figure shows the two BNN components ($\sigma_{\text{syst}}^{\text{BNN}}$, $\sigma_{\text{stat}}^{\text{BNN}}$) and the two RE ($\sigma_{\text{syst}}^{\text{RE}}$, $\sigma_{\text{stat}}^{\text{RE}}$) components separately. The results discussed in the following were achieved with a training dataset where N_{train} was sufficiently large to assume $\sigma_{\text{syst}} \gg \sigma_{\text{stat}} \Rightarrow \sigma_{\text{tot}} = \sigma_{\text{syst}}$.

5.3 Comparison of the learned uncertainties

The uncertainty predictions provided along with the learned topo-cluster calibration need to be analysed and interpreted to rule out that they are artefacts introduced by the machine learning model itself. Rather, they are supposed to reflect limitations in the data that lower the ability of the applied network model to accurately regress on the target, in addition to limitations introduced by the model itself.

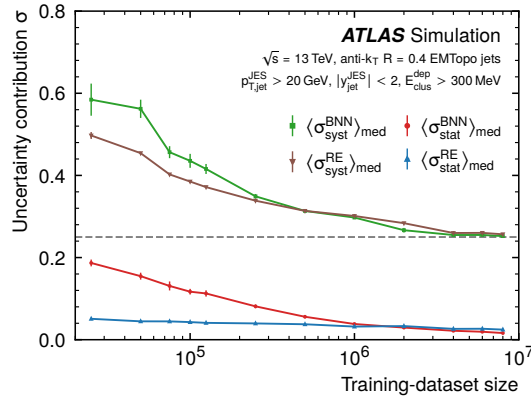


Figure 10: The dependence of the predicted uncertainty components on the training dataset size. Shown are the median uncertainty components discussed in Section 5 for training datasets of varying sizes and their specific asymptotic behaviours. The topo-clusters used here are collected into training datasets employing the selections shown in Section 3.1.1. The statistical uncertainties on the medians are indicated as vertical bars when exceeding the symbol size. The dashed horizontal line indicates the asymptotic limit of the systematic uncertainties.

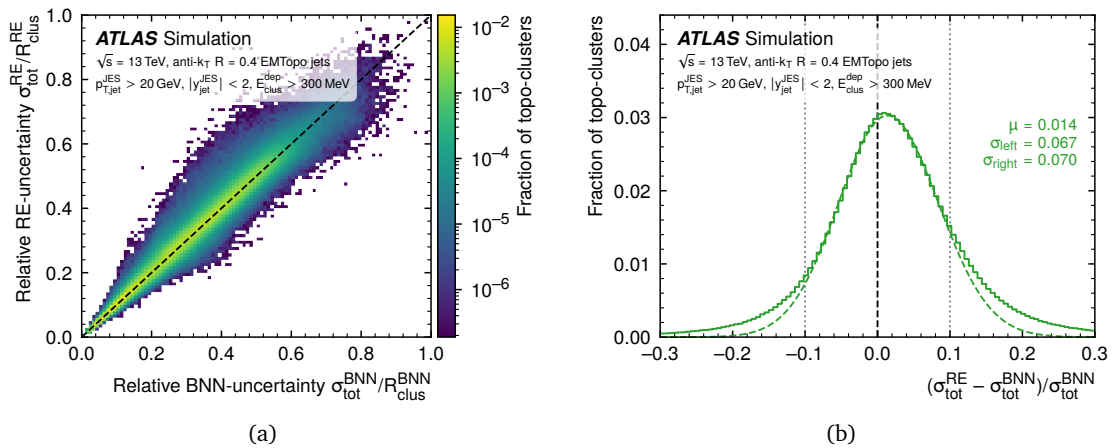


Figure 11: In (a), the relative total uncertainty prediction $\sigma_{\text{tot}}^{\text{RE}}/\mathcal{R}_{\text{clus}}^{\text{RE}}$ from the RE is shown as a function of the relative uncertainty $\sigma_{\text{tot}}^{\text{BNN}}/\mathcal{R}_{\text{clus}}^{\text{BNN}}$ from the BNN. The distribution of the relative difference between the two uncertainties $(\sigma_{\text{tot}}^{\text{RE}} - \sigma_{\text{tot}}^{\text{BNN}})/\sigma_{\text{tot}}^{\text{BNN}}$ is shown in (b). For illustration, the dashed (centre) line indicates the completely unbiased mean of the distribution, and the dotted lines indicate the 10% range around it. The dashed curve shows the result of a two-sided fit of Gaussian functions to the distribution, with a common mean μ and the independently fitted widths σ_{left} and σ_{right} . Both uncertainties are predicted for each topo-cluster in the test dataset. These clusters are collected from the simulated jets, as described in Section 3.1.1.

To address these concerns about specific network-introduced artefacts, the individual uncertainty predictions from the two different network models BNN and RE are directly compared cluster-by-cluster. Figure 11(a) shows the correlation between the relative learned uncertainties of the BNN ($\sigma_{\text{tot}}^{\text{BNN}}/\mathcal{R}_{\text{clus}}^{\text{BNN}}$) and the RE ($\sigma_{\text{tot}}^{\text{RE}}/\mathcal{R}_{\text{clus}}^{\text{RE}}$). As by construction the systematic uncertainties are not deterministic, a certain spread can be expected and is observed. Nevertheless, there is no sign of a bias in this correlation, the two predictions track each other well.

Figure 11(b) shows the distribution of the relative differences $(\sigma_{\text{tot}}^{\text{RE}} - \sigma_{\text{tot}}^{\text{BNN}})/\sigma_{\text{tot}}^{\text{BNN}}$ between the (absolute) BNN and the RE uncertainty predictions. A two-sided fit of Gaussian functions yields a near symmetric shape, indicating an agreement between the two estimates at the level of 10%. Differences to the fitted Gaussian functions in the tails between 10% and 30%, and -30% and -10% , are fairly identical and thus unbiased. They are expected for highly stochastic data like the response and the reconstructed features from topo-clusters with significant near-random contributions from the high levels of pile-up experienced in ATLAS at LHC Run 2 operations. This high level of comparability between the learned uncertainties shows that those predominately reflect features of the data, and not artefacts introduced by either network.

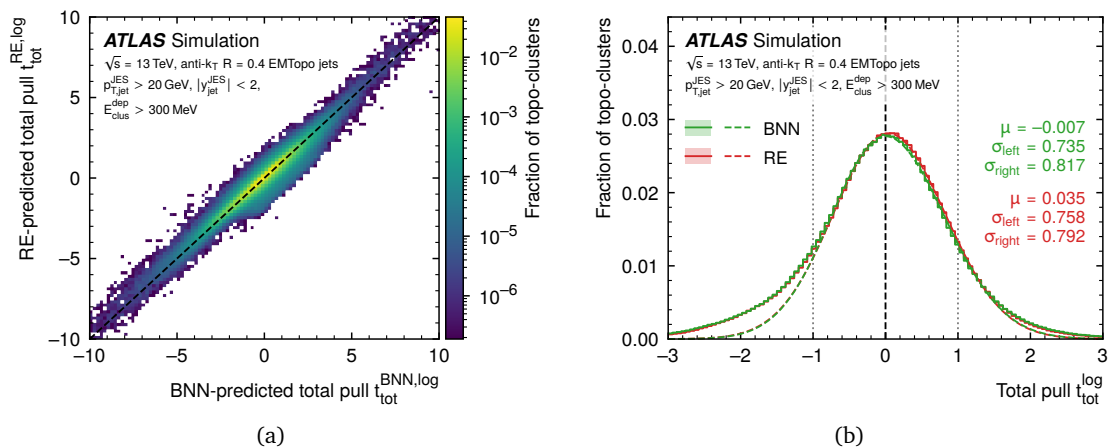


Figure 12: The observed pull from the RE training ($t_{\text{tot}}^{\text{RE}, \log}$) is shown as a function of the one from the BNN training ($t_{\text{tot}}^{\text{BNN}, \log}$) in (a). The inclusive distributions of $t_{\text{tot}}^{\text{BNN}, \log}$ and $t_{\text{tot}}^{\text{RE}, \log}$ are compare in (b). Both $t_{\text{tot}}^{\text{BNN}, \log}$ and $t_{\text{tot}}^{\text{RE}, \log}$ are calculated using Eq. 20, for a test dataset containing topo-clusters collected from full MC simulations, as described in Section 3.1.1. The dashed vertical line indicates the unbiased centre of the t_{tot}^{\log} distributions, while the dotted lines show the ± 1 range around it. The dashed curves indicate the two-sided fits of Gaussian functions for each network, with the common average μ and the independently fitted widths σ_{left} and σ_{right} . The corresponding fitted quantities are quoted.

5.4 Uncertainties and data spread

The interpretation from the previous section are augmented with the analysis of the predicted value of the uncertainty for any given topo-cluster. Such an analysis can be done using the pull $t_{\text{tot}}^{\text{BNN(RE)}}$,

$$t_{\text{tot}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}}) = \frac{\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}}) - \mathcal{R}_{\text{clus}}^{\text{EM}}}{\sigma_{\text{tot}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})}, \quad (19)$$

that measures the difference between the prediction $\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}$ from the BNN or RE and the target $\mathcal{R}_{\text{clus}}^{\text{EM}}$ on a scale defined by the respective predicted uncertainty $\sigma_{\text{tot}}^{\text{BNN(RE)}}$. Technically, the training is implemented such that it produces predictions for $\log_{10} \mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}$ for the target representation $\log_{10} \mathcal{R}_{\text{clus}}^{\text{EM}}$ and the uncertainty $\sigma_{\text{tot}}^{\text{BNN(RE),log}}$ at this transformed scale. The corresponding pull $t_{\text{tot}}^{\text{BNN(RE),log}}$ is related to the pull in the linear response space given in Eq. 19,¹²

$$\begin{aligned} t_{\text{tot}}^{\text{BNN(RE),log}}(\mathcal{X}_{\text{clus}}) &= \frac{\log_{10} \mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}}) - \log_{10} \mathcal{R}_{\text{clus}}^{\text{EM}}}{\sigma_{\text{tot}}^{\text{BNN(RE),log}}(\mathcal{X}_{\text{clus}})} \\ &\approx \ln 10 \cdot \log_{10} \frac{\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})}{\mathcal{R}_{\text{clus}}^{\text{EM}}} \cdot \frac{\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})}{\sigma_{\text{tot}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})} \\ &= \ln 10 \cdot \log_{10} \frac{\mathcal{R}_{\text{clus}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})}{\mathcal{R}_{\text{clus}}^{\text{EM}}} \cdot \left(t_{\text{tot}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}}) + \frac{\mathcal{R}_{\text{clus}}^{\text{EM}}}{\sigma_{\text{tot}}^{\text{BNN(RE)}}(\mathcal{X}_{\text{clus}})} \right). \end{aligned} \quad (20)$$

Figure 12 compares this pull in the logarithmic response space for the topo-cluster predictions from the BNN ($t_{\text{tot}}^{\text{BNN,log}}$) with the ones from the RE ($t_{\text{tot}}^{\text{RE,log}}$). Both $t_{\text{tot}}^{\text{BNN,log}}$ and $t_{\text{tot}}^{\text{RE,log}}$ are highly correlated, as shown in Figure 12(a). This indicates that both networks provide near identical pulls cluster-by-cluster, and is consistent with $\mathcal{R}_{\text{clus}}^{\text{BNN}} \approx \mathcal{R}_{\text{clus}}^{\text{RE}}$ for $\sigma_{\text{tot}}^{\text{RE}}/\mathcal{R}_{\text{clus}}^{\text{RE}} \approx \sigma_{\text{tot}}^{\text{BNN}}/\mathcal{R}_{\text{clus}}^{\text{BNN}}$ observed in the previous Section 5.3.

The pull distribution is expected to be close to a Gaussian distribution, with a central value $\langle t_{\text{tot}}^{\text{log}} \rangle = 0$ and a spread measured by its standard deviation of the order of unity, as motivated by the discussion in Appendix D.3. Figure 12(b) shows that both the $t_{\text{tot}}^{\text{BNN,log}}$ and the $t_{\text{tot}}^{\text{RE,log}}$ distributions follow Gaussian distributions in their bulk parts, with close to symmetric spreads determined by independent two-sided fits of σ_{left} and σ_{right} to each of these distributions. The asymmetry observed between the tails $t_{\text{tot}}^{\text{log}} < \sigma_{\text{left}}$, where some deviations from a Gaussian distribution are observed, and $t_{\text{tot}}^{\text{log}} > \sigma_{\text{right}}$, where the distribution more closely follows a Gaussian distribution, is an artefact introduced by the logarithmic transformation of the target and the respective predictions. The fitted spreads yield $\sigma_{\text{left}} \approx \sigma_{\text{right}} < 1$, indicating that both $\sigma_{\text{tot}}^{\text{BNN}}$ and $\sigma_{\text{tot}}^{\text{RE}}$ represent a conservative uncertainty estimate safely covering the residual difference between target and prediction for each topo-cluster. Considering these spreads together with central values fitted to be close to $t_{\text{tot}}^{\text{log}} = 0$, the conclusion from the pull distributions is that both $\sigma_{\text{tot}}^{\text{BNN}}$ and $\sigma_{\text{tot}}^{\text{RE}}$ are fairly unbiased and conservative uncertainties within expected value ranges.

5.5 Sources of uncertainties

The inclusive spectrum of $\sigma_{\text{tot}}^{\text{BNN}}$ learned by the BNN displayed in Figure 13 shows a rise of the population of topo-clusters with large uncertainties, deviating from the expected smoothly

¹²This relation uses $\sigma_{\text{tot}}^{\text{log}} \approx \frac{d \log_{10} R}{dR} \cdot \sigma_{\text{tot}} = \frac{1}{\ln 10 \cdot R} \cdot \sigma_{\text{tot}}$.

falling spectrum. The topo-clusters in the region $8.5 < \sigma_{\text{tot}}^{\text{BNN}} < 12.0$ around this emerging peak are investigated to search for possible reasons for its appearance. They show a large response $\mathcal{R}_{\text{clus}}^{\text{EM}} \gg 1$ and a small signal fraction in the electromagnetic calorimeter. Their centre-of-gravity is mostly located just in front of the endcap calorimeters, and within $1 \lesssim |y_{\text{clus}}^{\text{EM}}| \lesssim 1.4$. This points to the instrumentation of a specific transition range between ATLAS calorimeters generating the topo-cluster signals.

Figure 14(a) shows the topo-cluster centre-of-gravity projected into the $(|z_{\text{clus}}|, r_{\text{clus}})$ plane, where z_{clus} is its Cartesian coordinate along the central axis of the ATLAS detector and r_{clus} is its radial distance from this central axis, for all clusters in the test dataset. Selecting topo-clusters with $8.5 < \sigma_{\text{tot}}^{\text{BNN}} < 12.0$ from this sample yields the $(|z_{\text{clus}}|, r_{\text{clus}})$ -plane population depicted in Figure 14(b). The response and the features of these clusters are predominantly derived from signals in the Tile gap scintillator instrumenting the transition region between the central barrel and the endcap calorimeters. This detector is not a regular calorimeter in that it does not generate a signal corresponding to the energy loss generated in a (regular) sampling structure of active and inactive media in a calorimeter cell. Rather, it generates a signal reflecting the energy losses in the complex distribution of inactive material situated around it. This environment comprises cryostat walls, mechanical support structures and inner detector services. The energy deposits in these inactive materials are not considered in $\mathcal{R}_{\text{clus}}^{\text{EM}}$ defined in Eq. 4 of Section 3.2 for the topo-clusters located here, because they occur outside of calorimeter cells and are thus not accounted for in $E_{\text{clus}}^{\text{dep}}$.

As already discussed in Section 3.2.1, the actual signal $E_{\text{clus}}^{\text{EM}}$ of topo-clusters located in the Tile gap scintillators is predominantly generated by ionisations. Among the sources of these ionisations are secondary particles produced in the development of showers in the material around it. These showers are generated by particles emerging from the hard scatter and the in-time pile-up interactions. The latter can also provide a second and more direct contribution to the signal when directly crossing this transition region, without previous interaction. The BNN calibration network learns $\mathcal{R}_{\text{clus}}^{\text{EM}}$ from the feature set here, but the location of the topo-cluster centre-of-gravity suggests very little signal contributions from other calorimeters and thus to the feature reconstruction. In particular, features related to shower shapes are expected to be reconstructed with low precision, leading to a significantly reduced expressiveness that produces predictions with large uncertainties for topo-clusters in this region.

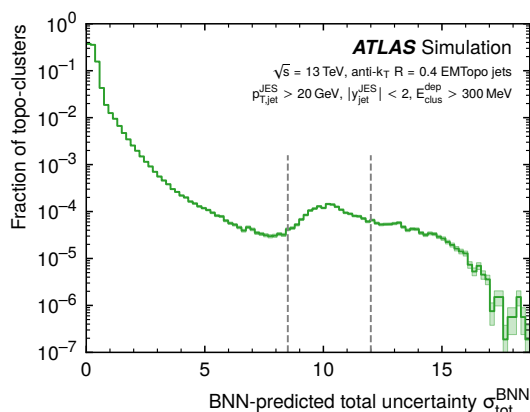


Figure 13: The inclusive spectrum of uncertainties $\sigma_{\text{tot}}^{\text{BNN}}$ learned by the BNN. The vertical lines indicate the region $8.5 < \sigma_{\text{tot}}^{\text{BNN}} < 12.0$ that is selected to analyse the source of the deviation from the expected smoothly falling spectrum. The uncertainties are predicted for topo-clusters from test dataset collected from the full MC simulations as described in Section 3.1.1.

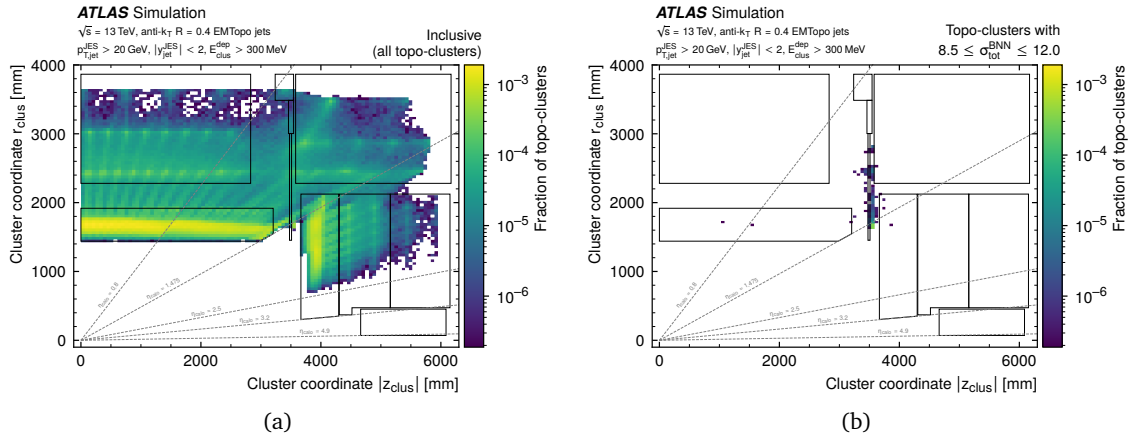


Figure 14: In (a), the centre-of-gravity location projected into the $(|z_{clus}|, r_{clus})$ plane is shown for all topo-clusters in the test dataset. The outlined structures indicate the instrumented area covered by the various ATLAS calorimeters in this plane. The lines depict selected pseudorapidities η_{calo} in the calorimeter pointing to transition regions. The lack of topo-clusters located beyond $|\eta_{calo}| \gtrsim 2.4$ is introduced by the jet selection described in Section 3.1.1. The location of the centre-of-gravity of topo-clusters with $8.5 < \sigma_{tot}^{BNN} < 12.0$ found in the same plane is shown in (b). These are predominantly inside, or in close proximity to, the thin Tile gap scintillators installed in the transition region between the central and the endcap calorimeters.

The fact that the BNN learns about these regional detector signal characteristics from the training data and predicts a large total uncertainty σ_{tot}^{BNN} that can be understood by relating it to the detector design, its signal formation, and its exposure to particles from the final state of the pp collision, suggests that a meaningful interpretation of σ_{tot}^{BNN} cluster-by-cluster is possible in, e.g., the context of topo-cluster signal accuracy and precision requirements in a given final-state analysis. Considering the previous discussions, σ_{tot}^{RE} is expected to reflect the same detector-induced issues.

6 Conclusions

The principal signal of the ATLAS calorimeters are clusters of topologically connected cell signals (topo-clusters). They are individually calibrated to measure the energy deposited in them using an uncertainty-aware Bayesian neural network (BNN). This network is designed to predict the response and the corresponding uncertainty for a topo-cluster as a function of a set of 15 features associated with it. These features are selected under consideration of expressiveness manifested in their sensitivity to the signal source at the location of the cluster, their dependence on the detector-specific signal formation characteristics, and the proton–proton collision environment characterised by pile-up.

The datasets collected for the BNN training, validation and testing comprise topo-clusters collected from calorimeter jets in full and detailed Monte Carlo simulations of multijet production in proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, with the ATLAS detector configurations in place at the time of the LHC Run 2 and with beam conditions generating levels of pile-up observed in this operational period (2015–2018).

The calibrated signal performance with respect to the deviation from signal linearity and the local relative energy resolution indicates significant improvements on most relevant

evaluation scales, if compared to the two standard calibrations in ATLAS, the basic electromagnetic topo-cluster calibration and the table-look-up-based hadronic calibration implemented in the LCW sequence, and to the calibration learned by a deep neural network (DNN) employing the same features as inputs as the BNN. In particular, the application of the BNN calibration extends near-perfect signal linearity to a previously underperforming region populated with topo-clusters with low-energy deposits from particles emerging from the hard scatter interaction. This calibration also improves the relative energy resolution at cluster level when compared with the resolution achieved with the DNN-based calibration, which already shows a significant improvement with respect to the standard calibrations.

The uncertainties predicted by the BNN consist of two components that in the case of the sufficiently large training dataset used here can be distinguished and respectively interpreted as statistical and systematic uncertainties. The BNN-learned uncertainties are confirmed using an alternative and independent calibration network featuring repulsive ensembles (RE), which predicts uncertainties that are cluster-by-cluster identical within the level of stochasticity in the dataset. An analysis of the pull distributions finds that the uncertainties from both networks follow a Gaussian distribution in the bulk, with some small deviations observed in particular in the low pull tails. The central values are found to be close to zero, indicating an unbiased pull. The spread of each pull distribution is less than the expected unity, which suggests that the learned uncertainties are conservative.

Large but appropriate uncertainties of the calibrated energy derived from the BNN response predictions are observed for topo-clusters located in the insufficiently instrumented transition region between the central and the endcap calorimeter modules. This shows promise for a potential applications in which the uncertainty of the calibrated topo-cluster signals can be used as evaluation tool for the signal quality and thus allow for selections based on accuracy (scale) and precision (resolution) requirements before the clusters enter the reconstruction of hadronic final-state objects like particles and jets, and the reconstruction of hadronic event shapes in general.

This study of the feasibility of the use of uncertainty-aware neural networks for the ATLAS calorimeter signal calibration in a realistically simulated proton–proton collision environment shows significant potential for a successful application in the context of physics analyses. The simulation-only results presented in this paper require investigations beyond these proof-of-principle studies, including the validation and performance evaluations for the application to experimental data at present day and at significantly increased future pile-up levels at the LHC. The BNN model and its inputs are deliberately designed to address these challenges by adapting the derived calibrations accordingly. This is mostly achieved by not using low-level data like calorimeter cell signals, but rather using features constructed from those as inputs. A potentially necessary retraining of the BNN in the case of a changing feature phase space due to changing operational conditions at the LHC is thus possible without the need for major signal reconstruction campaigns.

Acknowledgments

We thank CERN for the very successful operation of the LHC and its injectors, as well as the support staff at CERN and at our institutions worldwide without whom ATLAS could not be operated efficiently.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF/SFU (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and

large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [45].

Funding information We gratefully acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; D NRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ICHEP and Academy of Sciences and Humanities, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW, Poland; FCT, Portugal; MNE/IFA, Romania; MSTDI, Serbia; MSSR, Slovakia; ARIS and MVZI, Slovenia; DSI/NRF, South Africa; MICIU/AEI, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; NSTC, Taipei; TENMAK, Türkiye; STFC/UKRI, United Kingdom; DOE and NSF, United States of America.

Individual groups and members have received support from BCKDF, CANARIE, CRC and DRAC, Canada; CERN-CZ, FORTE and PRIMUS, Czech Republic; COST, ERC, ERDF, Horizon 2020, ICSC-NextGenerationEU and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and MINERVA, Israel; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

In addition, individual members wish to acknowledge support from Armenia: Yerevan Physics Institute (FAPERJ); CERN: European Organization for Nuclear Research (CERN DOCT); Chile: Agencia Nacional de Investigación y Desarrollo (FONDECYT 1230812, FONDECYT 1230987, FONDECYT 1240864); China: Chinese Ministry of Science and Technology (MOST-2023YFA1605700, MOST-2023YFA1609300), National Natural Science Foundation of China (NSFC - 12175119, NSFC 12275265, NSFC-12075060); Czech Republic: Czech Science Foundation (GACR - 24-11373S), Ministry of Education Youth and Sports (FORTE CZ.02.01.01/00/22_008/0004632), PRIMUS Research Programme (PRIMUS/21/SCI/017); EU: H2020 European Research Council (ERC - 101002463); European Union: European Research Council (ERC - 948254, ERC 101089007), European Union, Future Artificial Intelligence Research (FAIR-NextGenerationEU PE00000013), Italian Center for High Performance Computing, Big Data and Quantum Computing (ICSC, NextGenerationEU); France: Agence Nationale de la Recherche (ANR-20-CE31-0013, ANR-21-CE31-0013, ANR-21-CE31-0022, ANR-22-EDIR-0002); Germany: Baden-Württemberg Stiftung (BW Stiftung-Postdoc Eliteprogramme, BWST_IF2020-010 *Uncertainties – Teaching AI its Limits*), Deutsche Forschungsgemeinschaft (DFG - 469666862, DFG - CR 312/5-2, EXC 2181/1 – 390900948); Italy: Istituto Nazionale di Fisica Nucleare (ICSC, NextGenerationEU), Ministero dell’Università e della Ricerca (PRIN - 20223N7F8K - PNRR M4.C2.1.1); Japan: Japan Society for the Promotion of Science (JSPS KAKENHI JP22H01227, JSPS KAKENHI JP22H04944, JSPS KAKENHI JP22KK0227, JSPS KAKENHI JP23KK0245); Norway: Research Council of Norway (RCN-314472); Poland: Ministry of Science and Higher Education (IDUB AGH, POB8, D4 no 9722), Polish National Agency for Academic Exchange (PPN/PPO/2020/1/00002/U/00001), Polish National Science Centre (NCN 2021/42/E/ST2/00350, NCN OPUS 2023/51/B/ST2/02507, NCN OPUS nr 2022/47/B/ST2/03059, NCN UMO-2019/34/E/ST2/00393, NCN & H2020 MSCA 945339, UMO-2020/37/B/ST2/01043, UMO-2021/40/C/ST2/00187, UMO-2022/47/O/ST2/00148, UMO-2023/49/B/ST2/04085, UMO-2023/51/B/ST2/00920);

Spain: Generalitat Valenciana (Artemisa, FEDER, IDIFEDER/2018/048), Ministry of Science and Innovation (MCIN & NextGenEU PCI2022-135018-2, MICIN & FEDER PID2021-125273NB, RYC2019-028510-I, RYC2020-030254-I, RYC2021-031273-I, RYC2022-038164-I); Sweden: Carl Trygger Foundation (Carl Trygger Foundation CTS 22:2312), Swedish Research Council (Swedish Research Council 2023-04654, VR 2018-00482, VR 2022-03845, VR 2022-04683, VR 2023-03403, VR grant 2021-03651), Knut and Alice Wallenberg Foundation (KAW 2018.0458, KAW 2019.0447, KAW 2022.0358); Switzerland: Swiss National Science Foundation (SNSF - PCEFP2_194658); United Kingdom: Leverhulme Trust (Leverhulme Trust RPG-2020-004), Royal Society (NIF-R1-231091); United States of America: U.S. Department of Energy (ECA DE-AC02-76SF00515), Neubauer Family Foundation.

A Local hadronic calibration sequence

The standard local hadronic calibration for topo-clusters employs cell signal weighting for hadronic calibration (local cell weighting, LCW), right after the cluster classification estimating the likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$ of a cluster to be of electromagnetic origin, see sequence diagram shown in Figure 15. The electromagnetic and hadronic scale factors are applied at cell level such that the effective *cell signal weight* w_{cell}^{κ} at each step κ after the classification becomes

$$w_{\text{cell}}^{\kappa} = \mathcal{P}_{\text{clus}}^{\text{EM}} w_{\text{cell}}^{\text{em}-\kappa} + (1 - \mathcal{P}_{\text{clus}}^{\text{EM}}) w_{\text{cell}}^{\text{had}-\kappa}, \quad \text{with } \kappa \in \{\text{cal}, \text{ooc}, \text{dm}\}. \quad (\text{A.1})$$

Here $\mathcal{P}_{\text{clus}}^{\text{EM}}$ is reconstructed in the classification step, and the ‘em’ and ‘had’ prefixes indicate the algorithm from which the signal weights are derived. The calibrated topo-cluster energies after each LCW calibration step κ are reconstructed from sums of weighted cell energies, with weights

$$w_{\text{cell}}^{\kappa} w_{\text{cell}}^{\text{geo}} E_{\text{cell}},$$

applied to the cell signal E_{cell} at the EM scale. The geometrical weight $w_{E_{\text{cell}}}^{\text{geo}}$ is $w_{E_{\text{cell}}}^{\text{geo}} < 1$ for cells shared between two topo-clusters, and $w_{E_{\text{cell}}}^{\text{geo}} = 1$ for cells collected into one topo-cluster only.

The topo-cluster energy $E_{\text{clus}}^{\text{had}}$, which reflects the hadronic calibration applied to the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$ at step $\kappa = \text{cal}$ after classification, serves as a reference for comparisons with the machine-learned calibrations and is given by

$$E_{\text{clus}}^{\text{had}} = \sum_{i=1}^{N_{\text{clus}}^{\text{cell}}} w_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}} \left[\mathcal{P}_{\text{clus}}^{\text{EM}} w_{\text{cell},i}^{\text{had-em}} + (1 - \mathcal{P}_{\text{clus}}^{\text{EM}}) w_{\text{cell},i}^{\text{had-had}} \right] = \sum_{i=1}^{N_{\text{clus}}^{\text{cell}}} w_{\text{cell},i}^{\text{cal}} E_{\text{cell},i}^{\text{EM}}. \quad (\text{A.2})$$

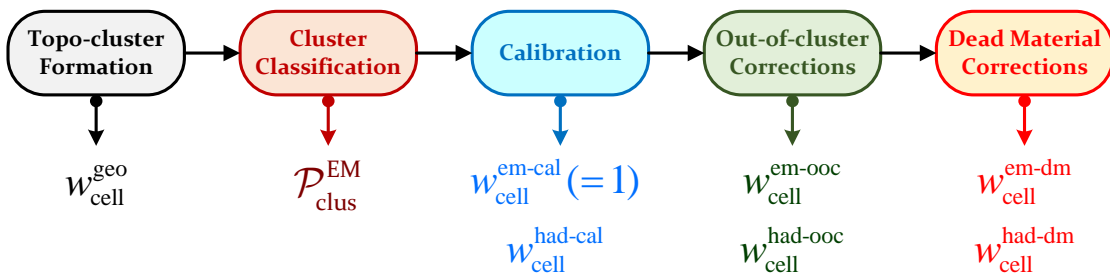


Figure 15: The LCW calibration sequence. Taken from Ref. [2].

Here $N_{\text{clus}}^{\text{cell}}$ is the total number of cells in the topo-cluster. Typical weights are $w_{\text{cell}}^{\text{em-cal}} = 1$, as E_{cell} is already at EM scale, and $w_{E_{\text{cell}}}^{\text{had-cal}} \geq 1$ to calibrate non-compensation ($e/h > 1$), both used to construct the effective cell signal weight $w_{\text{cell}}^{\text{cal}}$ at the calibration step. The electromagnetic and hadronic weights applied in the following correction steps for out-of-cluster and inactive material energy losses are typically larger than unity. As they depend on the nature of the underlying particle shower, they are typically also not identical for the respective electromagnetic and hadronic correction.

B Features

B.1 Feature construction

All features collected into the topo-cluster feature set $\mathcal{X}_{\text{clus}}$ defined in Eq. 3 and further described in Table 1 in Section 3.1.2, are documented in Ref. [2], except for the cluster compactness measure $p_{\text{T}}D$ and the variance $\text{Var}_{\text{clus}}(t_{\text{cell}})$ of the cell time distribution within the cluster.

The construction of $p_{\text{T}}D$ is inspired by a variable introduced in the context of quark-gluon jet tagging in Ref. [30]. While preserving the naming convention introduced in the original reference, the definition of $p_{\text{T}}D$ is slightly modified for computational efficiency when considering that clustered calorimeter cells are directionally in close proximity, by replacing the transverse momenta of the jet constituents in the original definition by the cell energy E_{cell} such that

$$p_{\text{T}}D = \frac{\sqrt{\sum_{i|E_{\text{cell},i}^{\text{EM}} > 0} (w_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}})^2}}{\sum_{i|E_{\text{cell},i}^{\text{EM}} > 0} w_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}}}. \quad (\text{B.1})$$

Only cells with energies $E_{\text{cell}} > 0$ are considered in the sums. By definition, $p_{\text{T}}D$ is numerically restricted to $0 < p_{\text{T}}D \leq 1$.

A compactness of $p_{\text{T}}D \rightarrow 1$ indicates that the energy $E_{\text{clus}}^{\text{EM}}$ of the topo-cluster is increasingly contained in only one cell. A compactness measure of $p_{\text{T}}D = 1$ exact is only possible for a single cell cluster, which is not expected to happen, due to the topo-cluster formation rules stating that at least all neighbours of a seed cell are collected [2]. Very small number of cells are possible for topo-clusters at the edges of the calorimeter, like at the outer circumference of the Tile, the back of the HEC, and the front faces of the electromagnetic calorimeters EMB and EMEC. A smaller value of $p_{\text{T}}D$ indicates a disperse distribution of the energy across the cells in the cluster.

The variance $\text{Var}_{\text{clus}}(t_{\text{cell}})$ of the cell time (t_{cell}) distribution is given by

$$\text{Var}_{\text{clus}}(t_{\text{cell}}) = \langle t_{\text{cell}}^2 \rangle - t_{\text{clus}}^2, \quad (\text{B.2})$$

where t_{cell} are the cell times and t_{clus} is the cluster time. The second moment $\langle t_{\text{cell}}^2 \rangle$ of the t_{cell} distribution for cells collected in the topo-cluster is defined by

$$\langle t_{\text{cell}}^2 \rangle = \frac{\sum_{i|\zeta_{\text{cell},i} > 2} (w_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}})^2 \cdot t_{\text{cell},i}^2}{\sum_{i|\zeta_{\text{cell},i} > 2} (w_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}})^2}, \quad (\text{B.3})$$

following the definition of t_{clus} , which is constructed as a first moment of the t_{cell} distribution [2]. Only cells with a signal significance (signal-over-noise) of $\zeta_{\text{cell}} > 2$ are considered when calculating t_{clus} and $\text{Var}_{\text{clus}}(t_{\text{cell}})$.

B.2 Relation of features to topo-cluster nature

The features collected into the feature set $\mathcal{X}_{\text{clus}}$ defined in Eq. 3 in Section 3.1.2 are evaluated with respect to the composition of the topo-cluster signal source $E_{\text{clus}}^{\text{dep}}$. For this, feature distributions for the three categories defined in Section 2.3.2, electromagnetic topo-clusters, hadronic topo-clusters and composite topo-clusters, are compared in Figures 16 and 17. Their relative contribution to the corresponding inclusive distributions is shown in same figures.

As the topo-clusters considered here are extracted from jets, the typical fraction of electromagnetic topo-clusters in the sample can be estimated. Using the canonical fraction of the jet energy carried by photons, about 25%, together with the average topo-cluster multiplicity for hadrons shown in Ref. [2] and the assumption that each photon in a jet generates one cluster, this fraction is about 10%. This estimate is only given as guideline, as it ignores all full or partial merging of shower signals into one topo-cluster and all shower signal losses introduced by the topological cell clustering algorithm and its configuration.

Other feature distributions with near-constant contributions from electromagnetic topo-clusters are $y_{\text{clus}}^{\text{EM}}$ shown in Figure 16(a), N_{PV} in Figure 17(e) and μ in Figure 17(f). Those also show approximately constant contributions from the other topo-cluster categories. This insensitivity to the cluster nature is expected, as neither $y_{\text{clus}}^{\text{EM}}$ nor N_{PV} or μ should show any effect on the composition of the energy deposits, which to first order only depend on the fragmentation of jets from the hard-scatter interactions. Slight variations may be introduced by the amount of inactive material in front of the active calorimeter volume that shows some variations with $y_{\text{clus}}^{\text{EM}}$. These features are kept in $\mathcal{X}_{\text{clus}}$ because they provide sensitivity to changes of the detector readout geometry, including the varying projection of the particle showers onto the calorimeter cell geometry and thus the reconstruction of most spatial, compactness and density features, and the pile-up environment.

Among the strongest indicators of the topo-cluster nature are $\zeta_{\text{clus}}^{\text{EM}}$ (Figure 16(b)), the location measures λ_{clus} (Figure 16(e)) and $|\vec{c}_{\text{clus}}|$ (Figure 16(f)), f_{emc} (Figure 16(g)), ρ_{clus} (Figure 16(h)), the dispersion and compactness measures $\langle m_{\text{long}}^2 \rangle$, $\langle m_{\text{lat}}^2 \rangle$ and $p_{\text{T}}D$ in Figures 17(a)–17(c). These figures all confirm expectations from electromagnetic and hadronic shower developments in the ATLAS calorimeters. For example, the contribution of electromagnetic topo-clusters to the f_{emc} distribution is generated by the photons in jets, nearly all of which are represented as a single cluster. They are mostly stopped in the electromagnetic calorimeters. This is in full agreement with the observation that electromagnetic topo-clusters are more isolated, as they contribute more than 25% of all clusters with $f_{\text{iso}} \gtrsim 0.8$, at an inclusive contribution of about 10%, as discussed above.

The larger contribution of electromagnetic topo-clusters to the population of higher-density clusters seen in Figure 16(h) is also well within the expectation of higher-density deposits in electromagnetic showers in the smaller volume cells of the electromagnetic calorimeters, with a higher signal yield when compared with hadronic showers. Also as expected, hadronic topo-clusters are dominantly contributing to the population of clusters deep in the calorimeter, as seen for the $|\vec{c}_{\text{clus}}|$ and λ_{clus} distributions in Figures 16(e) and 16(f). Compact topo-clusters with $p_{\text{T}}D \gtrsim 0.6$ are most likely of hadronic nature, as seen in Figure 17(c). These compact topo-clusters often have one dominant cell signal, which can occur in the course of the hadronic shower development in the hadronic calorimeters with their coarse spatial readout granularity and in the Tile calorimeter for signals from ionisations by charged hadrons. As already seen, electromagnetic topo-clusters are much more likely to appear in the highly granular electromagnetic calorimeters, where despite their high energy loss density signal sharing between cells is very likely, due to the relatively small cell size, compared to typical electromagnetic shower sizes.

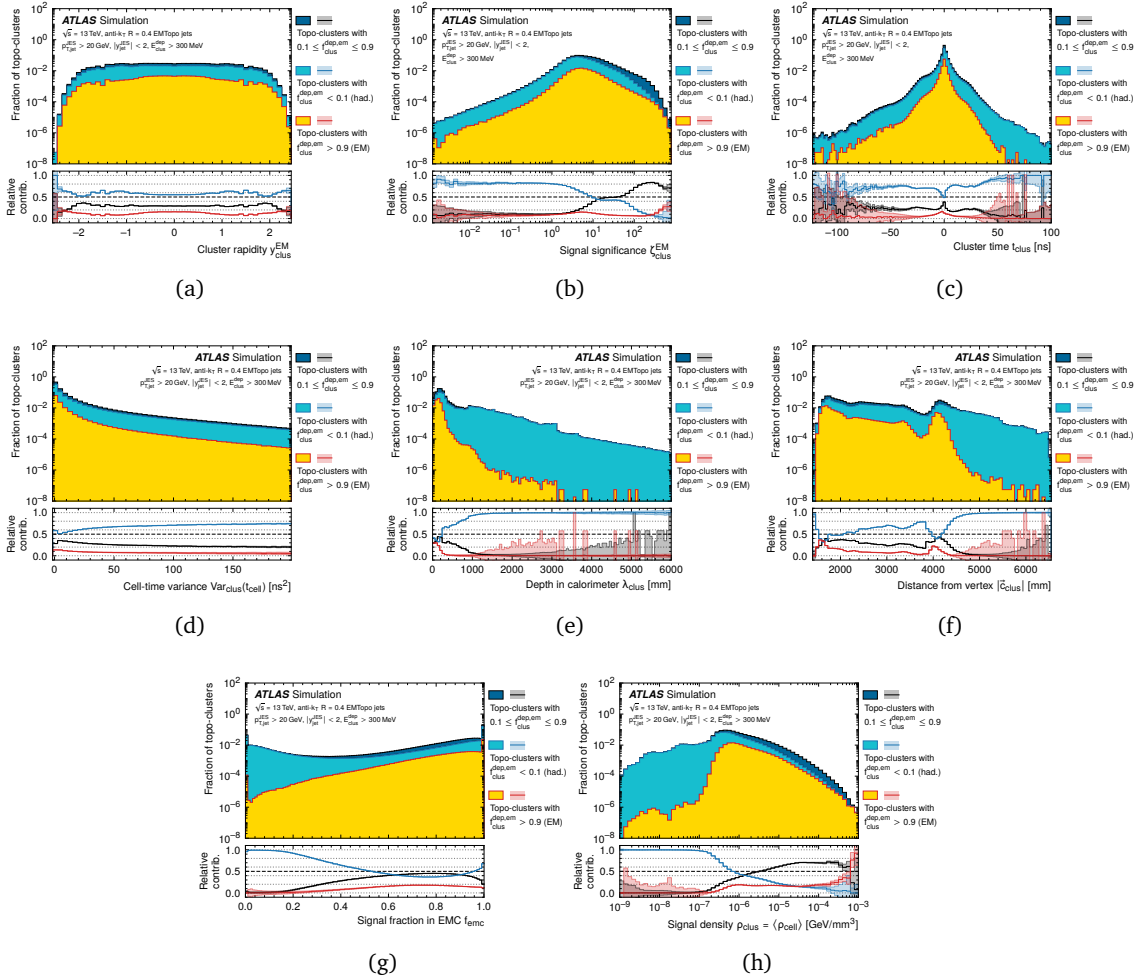


Figure 16: Distributions of some of the features given in Eq. 3 for topo-clusters classified as hadronic, electromagnetic or composite following the definitions given in Section 2.3.2. Distributions are shown for (a) the rapidity y_{clus}^{EM} , (b) the signal significance ζ_{clus}^{EM} , (c) the time t_{clus} , (d) the variance of the cell time distribution $\text{Var}_{clus}(t_{cell})$, (e) the distance from the calorimeter front-face λ_{clus} , (f) the distance from the nominal vertex $|\vec{c}_{clus}|$, (g) the fraction of energy in the electromagnetic calorimeter f_{emc} , and (h) the signal density ρ_{clus} . The fractional contributions from the three topo-cluster categories are shown in the respective lower panels of each figure. The shaded regions show the statistical uncertainties. More information about the definition of $\text{Var}_{clus}(t_{cell})$ is given in Appendix B.1. The other features are defined in Ref. [2]. An overview on all features is given in Table 1.

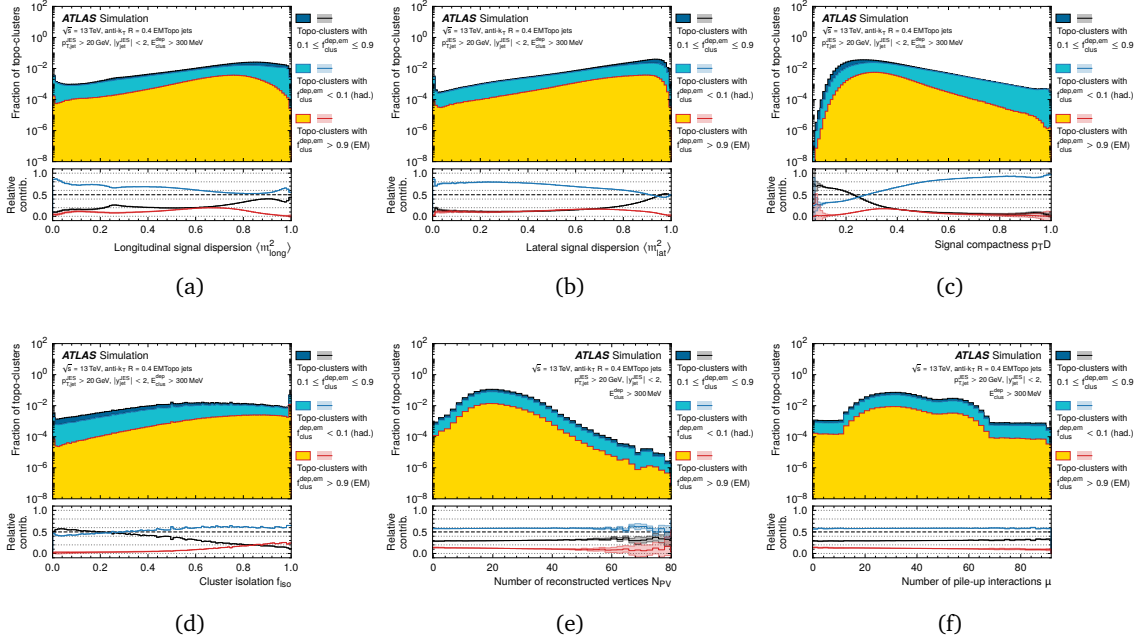


Figure 17: Distributions of some of the features given in Eq. 3 for topo-clusters classified as hadronic, electromagnetic or composite following the definitions given in Section 2.3.2. Distributions are shown for (a) the longitudinal energy dispersion $\langle m_{\text{long}}^2 \rangle$, (b) the lateral energy dispersion $\langle m_{\text{lat}}^2 \rangle$, (c) the compactness measure $p_T D$, (d) the isolation measure f_{iso} , (e) the number of reconstructed vertices N_{PV} , and (f) the number of pile-up interactions μ . The fractional contributions from the three topo-cluster categories are shown in the respective lower panels of each figure. The shaded regions show the statistical uncertainties. More information about the definition of $p_T D$ is given in Appendix B.1. The other features are defined in Ref. [2]. An overview on all features is given in Table 1.

C Network implementations and loss functions

C.1 Bayesian loss function

As discussed in Section 3.3, a Bayesian neural network (BNN) is employed to encode the response function $\mathcal{R}(\mathcal{X}_{\text{clus}}) = \mathcal{R}_{\text{clus}}^{\text{BNN}}(\mathcal{X}_{\text{clus}})$ and its uncertainty over the space of all features sets $\mathcal{X}_{\text{clus}}$ found for the topo-clusters in the training dataset. It is designed to learn $\mathcal{R} = \mathcal{R}(\mathcal{X}_{\text{clus}})$ as a probability distribution $p(\mathcal{R})$, with the mean

$$\langle \mathcal{R} \rangle = \int d\mathcal{R} \mathcal{R} p(\mathcal{R}). \quad (\text{C.1})$$

The network encodes the learned probability in weight configurations, conditional on the training data. The training can be described as constructing a variational approximation [46, 47], replacing $p(\theta|D_{\text{train}})$ introduced in Section 3.3 with a learned $q(\theta)$ such that

$$p(\mathcal{R}) = \int d\theta p(\mathcal{R}|\theta)p(\theta|D_{\text{train}}) \approx \int d\theta p(\mathcal{R}|\theta)q(\theta). \quad (\text{C.2})$$

The dataset D_{train} is defined over the feature space containing all possible feature sets $\mathcal{X}_{\text{clus}}$, following the nomenclature given in Eq. 11 of Section 3.3.2. To learn $q(\theta)$, the Kullback-Leibler

(KL) [38] divergence including the Bayesian transform $p(\theta|D_{\text{train}}) \rightarrow p(D_{\text{train}}|\theta)$

$$\begin{aligned}
 D_{\text{KL}}[q(\theta), p(\theta|D_{\text{train}})] &= \int d\theta q(\theta) \log \frac{q(\theta)}{p(\theta|D_{\text{train}})} \\
 &= \int d\theta q(\theta) \log \frac{q(\theta)p(D_{\text{train}})}{p_{\text{prior}}(\theta)p(D_{\text{train}}|\theta)} \\
 &\approx D_{\text{KL}}[q(\theta), p_{\text{prior}}(\theta)] - \underbrace{\int d\theta q(\theta) \log p(D_{\text{train}}|\theta)}_{\langle \log p(D_{\text{train}}|\theta) \rangle_{\theta \sim q(\theta)}}
 \end{aligned} \tag{C.3}$$

is minimized. The explicit form of $p(D_{\text{train}}|\theta)$ is given in Eq. 12 in Section 3.3.2.

The first term in Eq. C.3 arises from the prior $p_{\text{prior}}(\theta)$ and can be viewed as a weight regularisation. The second term samples from the negative logarithmic likelihood, where the sampling generalises the standard dropout in neural network training. The BNN loss is then defined as given in Eq. 13,

$$\mathcal{L}_{\text{BNN}} = D_{\text{KL}}[q(\theta), p_{\text{prior}}(\theta)] - \langle \log p(D_{\text{train}}|\theta) \rangle_{\theta \sim q(\theta)}. \tag{C.4}$$

A Gaussian distribution is used as a convenient prior $p_{\text{prior}}(\theta)$. Its width is a hyper-parameter that could be varied on a performance plateau to make the training more efficient [11]. If the weight distributions $q(\theta)$ are also chosen as Gaussian functions, the KL-divergence can be computed analytically. Provided the network is sufficiently deep, the assumption of independent Gaussian distributions for each network parameter should not affect its expressiveness.

C.2 Implementation of repulsive ensemble networks

The principle implementation of the repulsive ensemble (RE) network training is described in Ref. [3]. As already mentioned in Section 3.3.3, RE networks modify the update rule that minimises $-\log p(\theta|D_{\text{train}})$ by gradient descent. This is done by evaluating this update rule for an ensemble of N networks and introduce a repulsive term that forces the ensemble to spread out and cover the loss around the actual minimum. The repulsive term is designed to increase with the proximity of the network parameters θ of a given ensemble member to all other members. It is implemented by a Gaussian kernel $k(\theta, \theta_j)$, with $j = 1, \dots, N$, such that the update from training step t to training step $t + 1$ becomes

$$\theta^{t+1} = \theta^t + \alpha \nabla_{\theta^t} \log p(\theta|D_{\text{train}}) - \alpha \frac{\nabla_{\theta^t} \sum_{j=1}^N k(\theta^t, \theta_j^t)}{\sum_{i=1}^N k(\theta^t, \theta_i^t)}. \tag{C.5}$$

The first two terms represent the standard network training by gradient descent, with the learning rate α and a gradient ∇_{θ^t} reflecting the evolution of the network parameters θ^t at the training step $t \rightarrow t+1$. The additional kernel term runs over the N networks of an ensemble and defines repulsive ensembles.

Using properties of ordinary differential equations describing the network training as a time evolution and the unique stationary solution of the Fokker-Planck equation (see, e.g., Ref. [48]), it is found that this update rule leads to an ensemble of networks with parameters sampled from $\theta \sim p(\theta|D_{\text{train}})$.

To fit a calibration function for topo-clusters, and similar other regression tasks, the function encoded by the network is of interest, not its latent representation θ in weight space. For instance, two networks can encode the same function by permuting the weights, unaffected

by a repulsive force in weight space. The repulsive term therefore needs to act in the space of the actual network outputs $\mathcal{R}_{\text{clus}}^{\text{RE}}(\mathcal{X}_{\text{clus}}) = \mathcal{R}_{\theta}(\mathcal{X}_{\text{clus}})$ represented by θ . The update rule in this function space is given by

$$\frac{\mathcal{R}^{t+1} - \mathcal{R}^t}{\alpha} = \nabla_{\mathcal{R}^t} \log p(\mathcal{R} | D_{\text{train}}) - \frac{\sum_j \nabla_{\mathcal{R}^t} k(\mathcal{R}, \mathcal{R}_j)}{\sum_j k(\mathcal{R}, \mathcal{R}_j)}. \quad (\text{C.6})$$

It is formally similar to the one in the weight space given in Eq. C.5.

The network training is still performed in the weight space, so a translation of the function-space update rule into this space is needed,

$$\frac{\theta^{t+1} - \theta^t}{\alpha} = \nabla_{\theta^t} \log p(\theta^t | D_{\text{train}}) - \frac{\partial \mathcal{R}^t}{\partial \theta^t} \frac{\sum_j \nabla_{\mathcal{R}} k(\mathcal{R}_{\theta^t}, \mathcal{R}_{\theta_j^t})}{\sum_j k(\mathcal{R}_{\theta^t}, \mathcal{R}_{\theta_j^t})}. \quad (\text{C.7})$$

The repulsive kernel in the function space cannot be evaluated, so the function must be evaluated for a finite batch of topo-clusters represented by points $x = \mathcal{X}_{\text{clus}}$ in the feature space,

$$\frac{\theta^{t+1} - \theta^t}{\alpha} \approx \nabla_{\theta^t} \log p(\theta^t | D_{\text{train}}) - \frac{\sum_{j=1}^N \nabla_{\theta^t} k(\mathcal{R}_{\theta^t}(x), \mathcal{R}_{\theta_j^t}(x))}{\sum_{j=1}^N k(\mathcal{R}_{\theta^t}(x), \mathcal{R}_{\theta_j^t}(x))}. \quad (\text{C.8})$$

Here $\mathcal{R}_{\theta^t}(x)$ in Eqs. C.7 and C.8 is understood as evaluating the function \mathcal{R} for members x of D_{train} at training step t .

C.3 Repulsive ensemble loss function

The modified update rule of Eq. C.8 can be turned into a loss function for the RE training. Using Bayes' theorem and neglecting the evidence $p(D_{\text{train}})$, the likelihood loss measure $p(\theta | D_{\text{train}})$ can be written as

$$\log p(\theta | D_{\text{train}}) = \log p(D_{\text{train}} | \theta) - \underbrace{\frac{\theta^2}{2\sigma^2}}_{\text{exponential of } p_{\text{prior}}(\theta)} + \text{const}. \quad (\text{C.9})$$

The prior $p_{\text{prior}}(\theta)$ is chosen to be Gaussian.

For a training dataset of size N_{train} , split into batches of size N_{batch} , $\mathcal{R}_{\theta^t}(x)$ is evaluated as the function of members $x_1, \dots, x_{N_{\text{batch}}}$ for all batches of the training dataset D_{train} . The update rule from Eq. C.8 can be rewritten using Eq. C.9,

$$\frac{\theta^{t+1} - \theta^t}{\alpha} \approx \nabla_{\theta^t} \frac{N_{\text{train}}}{N_{\text{batch}}} \sum_{b=1}^{N_{\text{batch}}} \log p(x_b | \theta) - \frac{\sum_{j=1}^N \nabla_{\theta^t} k(\mathcal{R}_{\theta^t}(x), \mathcal{R}_{\theta_j^t}(x))}{\sum_{j=1}^N k(\mathcal{R}_{\theta^t}(x), \mathcal{R}_{\theta_j^t}(x))} - \nabla_{\theta^t} \frac{\theta^2}{2\sigma^2}. \quad (\text{C.10})$$

Based on this rule, the loss function is constructed by inverting the sign in front of the gradient in Eq. C.10 and divide by the training dataset size N_{train} to remove the scaling with N_{train} . In addition, not all occurrences of θ are inside the gradient, so a *stop-gradient operation* is introduced, denoted by $\overline{\mathcal{R}}_{\theta_j}(x)$. With this, the RE loss function is given by

$$\mathcal{L}_{\text{RE}} = \sum_{i=1}^N \left[\underbrace{-\frac{1}{N_{\text{batch}}} \sum_{b=1}^{N_{\text{batch}}} \log p(x_b | \theta_i)}_{(1) \text{ negative logarithmic likelihood averaged over batch}} + \overbrace{\frac{1}{N_{\text{train}}} \frac{\sum_{j=1}^N k(\mathcal{R}_{\theta_i}(x), \overline{\mathcal{R}}_{\theta_j}(x))}{\sum_{j=1}^N k(\overline{\mathcal{R}}_{\theta_i}(x), \overline{\mathcal{R}}_{\theta_j}(x))}}^{(2) \text{ repulsive term in function space}} + \underbrace{\frac{1}{N_{\text{train}}} \frac{\theta_i^2}{2\sigma^2}}_{(3) \text{ normalised exponent of } p_{\text{prior}}(\theta_i)} \right]. \quad (\text{C.11})$$

Here the numbers (1), (2) and (3) refer to the discussion of the corresponding terms in Eq. 14 of Section 3.3.4. The normalised exponent of the prior implements a ridge (L_2) regularisation (see, for example, Ref. [39]). It controls possible over-fitting and focusses the fit on regions of relevant data.

A typical choice for the repulsive term is a Gaussian kernel, with a width given by the *median heuristic*¹³ [49] between the N ensemble members. It is divided by the training dataset size N_{train} to remove any scaling with N_{train} .

D Evaluation metrics

D.1 Performance

For the evaluation of the performance of the machine-learned topo-cluster calibration, distributions of test variables like the prediction power $\Delta_{\mathcal{R}}$ defined in Eq. 15 in Section 4.1.1 and the deviation from signal linearity Δ_E defined in Eq. 16 in Section 4.1.2 are characterised by a measure of the central tendency and a corresponding measure of their spread. The median value of the distribution is used to express the central tendency, because it is an unbiased measure that does not require any assumption on the shape of the distribution (like the mode does) and it shows the least dependence on the sample size, compared with the statistical average of an asymmetric distribution with developing tails.

The spread is defined by the *inter-quantile range* $Q_{f=68\%}^w$,

$$Q_{f=68\%}^w = \min(x_k - x_i) = x_q^{\text{right}} - x_q^{\text{left}}, \quad \text{such that} \quad \frac{\int_{x_q^{\text{left}}}^{x_q^{\text{right}}} dx p(x)}{\int_{-\infty}^{+\infty} dx p(x)} = 0.68, \quad (\text{D.1})$$

which for any distribution $p(x)$ is the smallest of all x value ranges that contains 68% of the collected statistics. This range corresponds to the $[-1\sigma, +1\sigma]$ range around the mean of a Gaussian distribution, and is symmetric around any measure of central tendency in case of symmetric distributions. Here x_q^{left} and x_q^{right} are the respective lower and upper limit of the x -value range defining the *inter-quantile region* $\text{IQR}_{f=68\%}$,

$$\text{IQR}_{f=68\%} = [x_q^{\text{left}}, x_q^{\text{right}}] = [x_q^{\text{left}}, x_q^{\text{left}} + Q_{f=68\%}^w]. \quad (\text{D.2})$$

With the choices of distribution descriptors given above, a measure for the relative energy resolution σ_{rel} can be constructed from the Δ_E distributions that employs both its median $\langle \Delta_E \rangle_{\text{med}}$ and its spread $Q_{f=68\%}^w(\Delta_E)$,

$$\sigma_{\text{rel}} = \frac{Q_{f=68\%}^w(\Delta_E)}{2(\langle \Delta_E \rangle_{\text{med}} + 1)}. \quad (\text{D.3})$$

If Δ_E is distributed following a Gaussian, $Q_{f=68\%}^w(\Delta_E) = 2\sigma(\Delta_E)$ such that

$$\sigma_{\text{rel}} = \frac{\sigma(\Delta_E)}{\langle \Delta_E \rangle + 1} = \frac{\sigma(E_{\text{clus}}^{\text{BNN(RE)}})}{\langle E_{\text{clus}}^{\text{BNN(RE)}} \rangle} \approx \frac{\sigma(E_{\text{clus}}^{\text{BNN(RE)}})}{\langle E_{\text{clus}}^{\text{dep}} \rangle},$$

is the relative energy resolution for calorimeters. Equation D.3 can therefore be considered as an appropriate representation of this resolution.

¹³This is the median of all pairwise distances between θ_i and θ_j for all N ensemble members ($i = 1, \dots, N$ and $j = 1, \dots, N$).

D.2 Uncertainties

The learned uncertainties from the BNN are derived by sampling the network parameters θ and calculating the variance by

$$\begin{aligned} \langle \mathcal{R} \rangle &= \int d\theta q(\theta) \langle \mathcal{R} \rangle_\theta, \quad \text{with} \quad \langle \mathcal{R} \rangle_\theta = \int d\mathcal{R} \mathcal{R} p(\mathcal{R}|\theta), \\ \sigma_{\text{tot}}^2 &= \int d\theta q(\theta) [\langle \mathcal{R}^2 \rangle_\theta - \langle \mathcal{R} \rangle_\theta^2 + (\langle \mathcal{R} \rangle_\theta - \langle \mathcal{R} \rangle)^2] \equiv \sigma_{\text{syst}}^2 + \sigma_{\text{stat}}^2. \end{aligned} \quad (\text{D.4})$$

Here $\langle \mathcal{R} \rangle$ is the expectation value for the prediction \mathcal{R} considering the probability $p(\mathcal{R}|\theta)$ that the actual θ represents \mathcal{R} , and σ_{tot}^2 is the variance associated with this expectation.¹⁴

The total uncertainty σ_{tot}^2 factorises into two terms. The first term in Eq. D.4 vanishes in the limit of arbitrarily well-known data, $p(\mathcal{R}|\theta) \rightarrow \delta(\mathcal{R}_{\text{clus}}^{\text{BNN}}(\theta) - \mathcal{R}_{\text{clus}}^{\text{EM}})$,

$$\sigma_{\text{syst}}^2 = \int d\theta q(\theta) [\langle \mathcal{R}^2 \rangle_\theta - \langle \mathcal{R} \rangle_\theta^2]. \quad (\text{D.5})$$

This uncertainty component approaches a plateau for large training data sets, and is therefore considered as a systematic uncertainty that cannot be reduced by increasing dataset sizes or the number of training attempts (epochs). It is predicted for each individual topo-cluster that is calibrated using $\mathcal{R}_{\text{clus}}^{\text{BNN}}$.

The second contribution is the θ -sampled variance

$$\sigma_{\text{stat}}^2 = \int d\theta q(\theta) [\langle \mathcal{R} \rangle_\theta - \langle \mathcal{R} \rangle]^2. \quad (\text{D.6})$$

It vanishes in the limit of perfect training, $q(\theta) \rightarrow \delta(\theta - \theta_0)$. This limit can be approached for large training datasets. This uncertainty component is therefore referred to as statistical.

Contrary to the BNN, where the predictions are sampled in the variational weight space defined by θ to determine the uncertainties, the RE approach samples the predictions in function space, as briefly introduced in Appendix C.3. Also in this case, and for sufficient large training datasets like the one used for the topo-cluster calibration, statistical and systematic uncertainty components can be distinguished and compared to the ones from the BNN.

D.3 Expectations for the pull distribution

The pull distributions shown in Figure 12(b) can be formally understood, and expectations for its width can be derived from the following considerations.

Generally, the pull of a learned distribution is defined as

$$t(x) = \frac{f_\theta(x) - f(x)}{\sigma(x)}, \quad (\text{D.7})$$

where $f_\theta(x)$ is the learned function with parameters θ , $f(x)$ the true function encoded in the training data, and $\sigma(x)$ is the learned uncertainty. For the calibration training discussed in this paper, $f_\theta(x) = \mathcal{R}_\theta(\mathcal{X}_{\text{clus}})$ for a learned set of network parameters θ and topo-cluster feature sets $\mathcal{X}_{\text{clus}}$, and $f(x) = \mathcal{R}_{\text{clus}}^{\text{EM}}(\mathcal{X}_{\text{clus}})$.

Ideally, $\sigma(x)$ captures the absolute value of the deviation of the learned function from the target exactly and for any x -value,

$$f_\theta(x) = f(x) \pm \sigma(x) \quad \Rightarrow \quad t(x) = \pm 1. \quad (\text{D.8})$$

¹⁴ $\langle \mathcal{R} \rangle_\theta^2$ is defined in analogy to $\langle \mathcal{R} \rangle_\theta$.

Alternatively, $\sigma(x)$ can capture the maximum absolute value of the deviation, but the actual deviation of $f_\theta(x)$ from $f(x)$ might be smaller. This actual deviation might also follow a flat distribution within the limits of $t(x) = \pm 1$,

$$\begin{aligned} p(f_\theta(x)) &= \begin{cases} \text{const.}, & f_\theta(x) \in [f(x) - \sigma(x), f(x) + \sigma(x)], \\ 0, & f_\theta(x) \notin [f(x) - \sigma(x), f(x) + \sigma(x)], \end{cases} \\ \Rightarrow p(t(x)) &= \begin{cases} \text{const.}, & t \in [-1, 1], \\ 0, & t \notin [-1, 1]. \end{cases} \end{aligned} \quad (\text{D.9})$$

For a stochastic system like the one represented by the topo-clusters, the learned values $f_\theta(x)$ are expected have a mean $\langle f_\theta(x) \rangle = f(x)$, but with a Gaussian smearing $\mathcal{N}(f(x), \sigma(x))$ with x -dependent width $\sigma(x)$ introduced by the (mostly pile-up) noise,

$$\begin{aligned} p(f_\theta(x)) &= \mathcal{N}(f(x), \sigma(x)) \propto \exp\left(-\frac{[f_\theta(x) - f(x)]^2}{2\sigma(x)^2}\right) \\ \Rightarrow p(t(x)) &= \mathcal{N}(0, 1). \end{aligned} \quad (\text{D.10})$$

This shows that for a stochastic smearing the pull follows a standard Gaussian distribution.

E Calibration performance

E.1 Performance in feature space

The deviation from signal linearity is evaluated before and after local calibrations comprising the hadronic calibration from LCW, the BNN-derived and the DNN-derived calibrations. Before local calibration, the topo-cluster signal $E_{\text{clus}}^{\text{EM}}$ is provided on the basic EM scale. As discussed in Section 4.1.2, $E_{\text{clus}}^{\text{EM}}$ is considered as uncalibrated and cannot be expected to have a direct proportionality with the deposited energy $E_{\text{clus}}^{\text{dep}}$. The topo-cluster energy at the hadronic scale from LCW $E_{\text{clus}}^{\text{LCW}}$ is expected to represent a calibrated signal, thus it is expected to be calibrated. Both of these scales do not use any of the features presented here as inputs to their respective signal reconstruction. Both topo-cluster energies calibrated using the BNN ($E_{\text{clus}}^{\text{BNN}}$) or the DNN ($E_{\text{clus}}^{\text{DNN}}$) predictions are expected to show a more linear response, as all features are used as inputs to their respective networks.

Figure 18 summarises the deviation from linearity for all four signal scales as a function of features only used for the BNN and DNN predictions.

E.2 Local relative energy resolution

Figure 19 shows the median deviation from linearity $\langle \Delta_E \rangle_{\text{med}}$ for the topo-cluster energy scales after applying the calibrations learned by the BNN and the DNN, as a function of the energy $E_{\text{clus}}^{\text{dep}}$ deposited in the cluster. The slices (a), (b) and (c) show individual distributions of Δ_E in selected bins of $E_{\text{clus}}^{\text{dep}}$, with the clear sign of significantly suppressed tails for the BNN-derived calibration at lower $E_{\text{clus}}^{\text{dep}}$. At higher $E_{\text{clus}}^{\text{dep}} \gtrsim 600$ MeV both predictions show increasingly similar distributions, with $\langle \Delta_E^{\text{BNN}} \rangle_{\text{med}} \approx 0$ and $\langle \Delta_E^{\text{DNN}} \rangle_{\text{med}} < 0$ for $E_{\text{clus}}^{\text{dep}} \gtrsim 80$ GeV indicating a more accurate median calibration by the trained BNN.

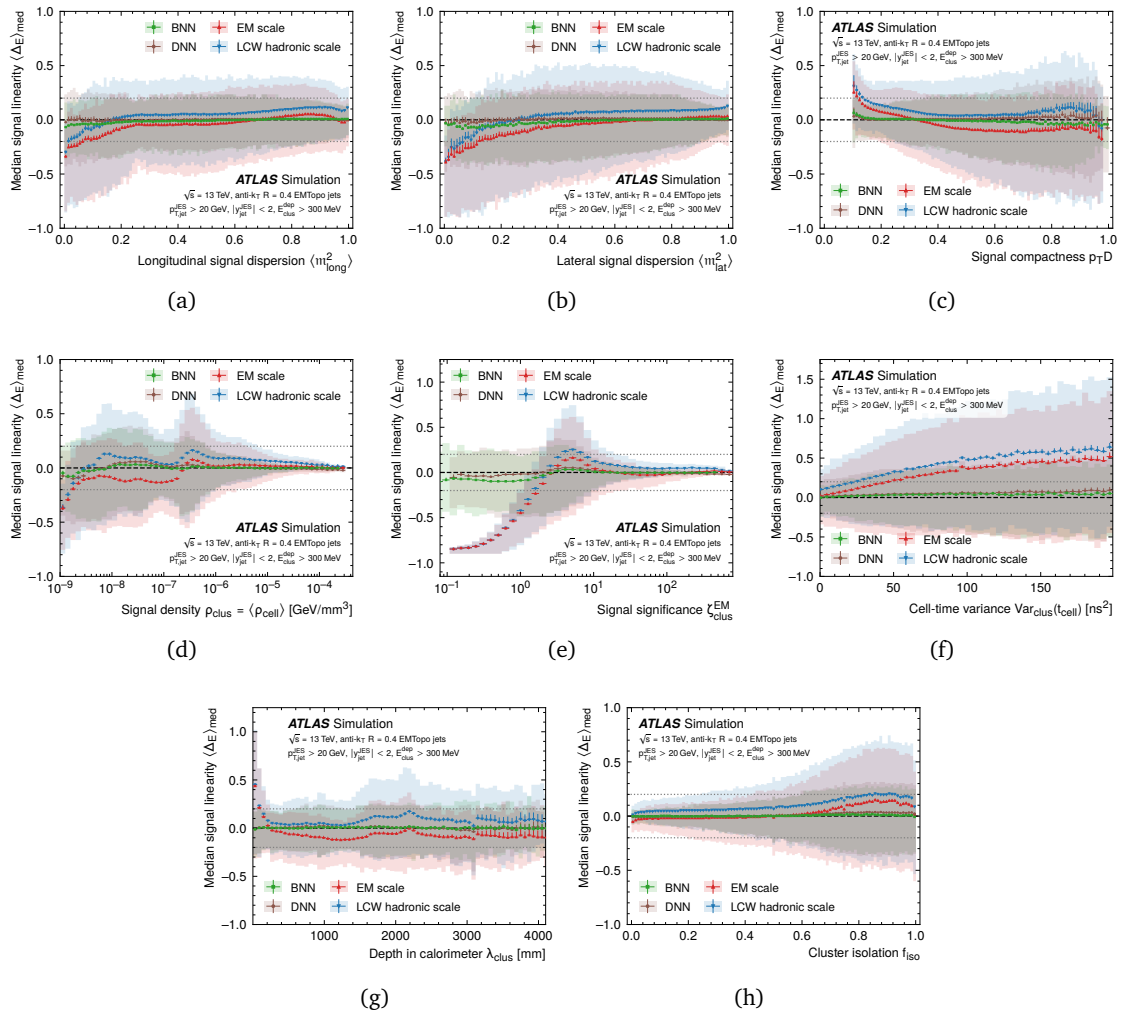
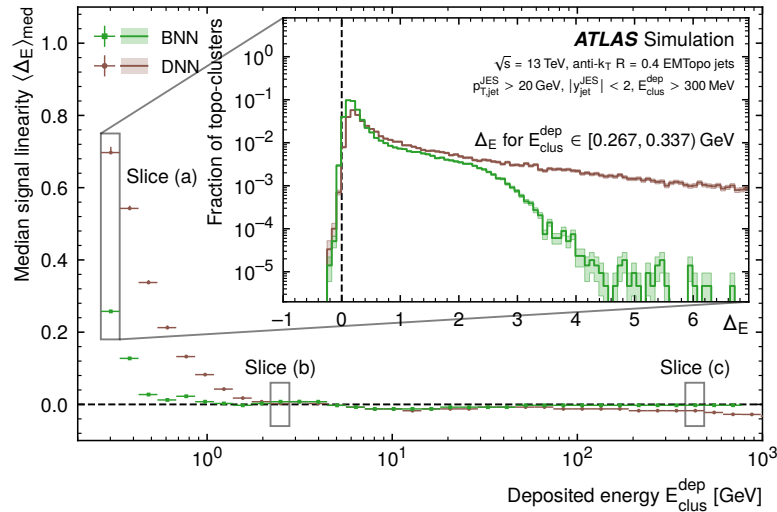
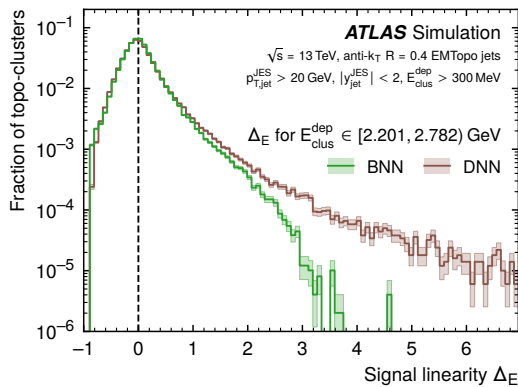


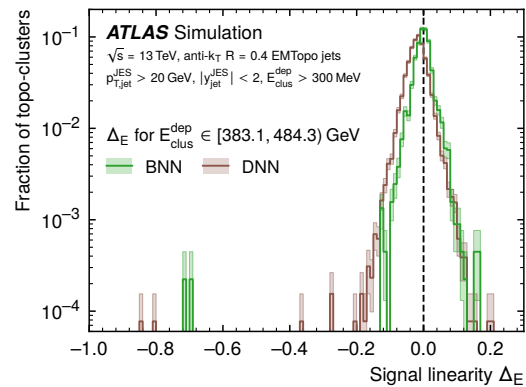
Figure 18: The median deviation from signal linearity for the topo-cluster at the basic EM scale ($\langle \Delta E^{\text{EM}} \rangle_{\text{med}}$), the hadronic calibration from LCW ($\langle \Delta E^{\text{had}} \rangle_{\text{med}}$), and at the BNN-learned ($\langle \Delta E^{\text{BNN}} \rangle_{\text{med}}$) and the DNN-learned ($\langle \Delta E^{\text{DNN}} \rangle_{\text{med}}$) calibrated energy scales, evaluated as a function of (a) the longitudinal signal dispersion measure $\langle m_{\text{long}}^2 \rangle$, (b) the lateral signal dispersion measure $\langle m_{\text{lat}}^2 \rangle$, (c) the signal compactness p_{TD} (see Appendix B.1), (d) the cluster signal density measure ρ_{clus} , (e) the cluster signal significance $\zeta_{\text{clus}}^{\text{EM}}$, (f) the second moment of the energy-squared-weighted cell time distribution $\text{Var}_{\text{clus}}(t_{\text{cell}})$ (see Appendix B.1), (g) the distance of the cluster centre-of-gravity from the calorimeter front face λ_{clus} , and (h) the cluster isolation measure f_{iso} . The topo-clusters of the test sample used to produce these results are extracted as described in Section 3.1.1.



(a)



(b)



(c)

Figure 19: In (a), the median deviation from linearity $\langle \Delta_E \rangle_{\text{med}}$ is shown as a function of the energy deposited in a topo-cluster $E_{\text{clus}}^{\text{dep}}$ for both the BNN- and the DNN-derived calibrations. The inset shows the shapes of the Δ_E^{BNN} and the Δ_E^{DNN} distributions in Slice (a), for $E_{\text{clus}}^{\text{dep}}$ around 300 MeV. The corresponding distributions for Slice (b) with $E_{\text{clus}}^{\text{dep}}$ around 2.5 GeV and Slice (c) with $E_{\text{clus}}^{\text{dep}}$ around 430 GeV are shown in (b) and (c), respectively. The shaded areas indicate the statistical uncertainties. The topo-clusters from the test dataset used to produce these results are extracted as described in Section 3.1.1.

References

- [1] ATLAS collaboration, *The ATLAS experiment at the CERN Large Hadron Collider*, J. Instrum. **3**, S08003 (2008), doi:[10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [2] ATLAS collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77**, 490 (2017), doi:[10.1140/epjc/s10052-017-5004-5](https://doi.org/10.1140/epjc/s10052-017-5004-5) [preprint doi:[10.48550/arXiv.1603.02934](https://arxiv.org/abs/1603.02934)].
- [3] T. Plehn, A. Butter, B. Dillon, T. Heimel, C. Krause and R. Winterhalder, *Modern machine learning for LHC physicists*, (arXiv preprint) doi:[10.48550/arXiv.2211.01421](https://doi.org/10.48550/arXiv.2211.01421).
- [4] ATLAS collaboration, *The application of neural networks for the calibration of topological cell clusters in the ATLAS calorimeters*, Tech. Rep. ATL-PHYS-PUB-2023-019, CERN, Geneva, Switzerland (2023), <https://cds.cern.ch/record/2866591>.
- [5] ATLAS collaboration, *Deep learning for pion identification and energy calibration with the ATLAS detector*, Tech. Rep. ATL-PHYS-PUB-2020-018, CERN, Geneva, Switzerland (2020), <https://cds.cern.ch/record/2724632>.
- [6] ATLAS collaboration, *Point cloud deep learning methods for pion reconstruction in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2022-040, CERN, Geneva, Switzerland (2022), <https://cds.cern.ch/record/2825379>.
- [7] D. J. C. Mackay, *Probable networks and plausible predictions — A review of practical Bayesian methods for supervised neural networks*, Netw.: Comput. Neural Syst. **6**, 469 (1995), doi:[10.1088/0954-898x_6_3_011](https://doi.org/10.1088/0954-898x_6_3_011).
- [8] R. M. Neal, *Bayesian learning for neural networks*, Springer, New York, USA, ISBN 9780387947242 (1996), doi:[10.1007/978-1-4612-0745-0](https://doi.org/10.1007/978-1-4612-0745-0).
- [9] Y. Gal, *Uncertainty in deep learning*, PhD thesis, University of Cambridge, Cambridge, UK (2016), <https://www.cs.ox.ac.uk/people/yarin.gal/website/thesis/thesis.pdf>.
- [10] A. Kendall and Y. Gal, *What uncertainties do we need in Bayesian deep learning for computer vision?*, (arXiv preprint) doi:[10.48550/arXiv.1703.04977](https://doi.org/10.48550/arXiv.1703.04977).
- [11] S. Bollweg, M. Haussmann, G. Kasieczka, M. Luchmann, T. Plehn and J. Thompson, *Deep-learning jets with uncertainties and more*, SciPost Phys. **8**, 006 (2020), doi:[10.21468/SciPostPhys.8.1.006](https://doi.org/10.21468/SciPostPhys.8.1.006) [preprint doi:[10.48550/arXiv.1904.10004](https://arxiv.org/abs/1904.10004)].
- [12] G. Kasieczka, M. Luchmann, F. Otterpohl and T. Plehn, *Per-object systematics using deep-learned calibration*, SciPost Phys. **9**, 089 (2020), doi:[10.21468/SciPostPhys.9.6.089](https://doi.org/10.21468/SciPostPhys.9.6.089) [preprint doi:[10.48550/arXiv.2003.11099](https://arxiv.org/abs/2003.11099)].
- [13] S. Badger, A. Butter, M. Luchmann, S. Pitz and T. Plehn, *Loop amplitudes from precision networks*, SciPost Phys. Core **6**, 034 (2023), doi:[10.21468/SciPostPhysCore.6.2.034](https://doi.org/10.21468/SciPostPhysCore.6.2.034) [preprint doi:[10.48550/arXiv.2206.14831](https://arxiv.org/abs/2206.14831)].
- [14] ATLAS collaboration, *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81**, 689 (2021), doi:[10.1140/epjc/s10052-021-09402-3](https://doi.org/10.1140/epjc/s10052-021-09402-3) [preprint doi:[10.48550/arXiv.2007.02645](https://arxiv.org/abs/2007.02645)].
- [15] F. D'Angelo and V. Fortuin, *Repulsive deep ensembles are Bayesian*, (arXiv preprint) doi:[10.48550/arXiv.2106.11642](https://doi.org/10.48550/arXiv.2106.11642).

- [16] L. Röver, B. M. Schäfer and T. Plehn, *PINNFerring the Hubble function with uncertainties*, (arXiv preprint) doi:[10.48550/arXiv.2403.13899](https://doi.org/10.48550/arXiv.2403.13899).
- [17] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **83**, 982 (2023), doi:[10.1140/epjc/s10052-023-11747-w](https://doi.org/10.1140/epjc/s10052-023-11747-w) [preprint doi:[10.48550/arXiv.2212.09379](https://doi.org/10.48550/arXiv.2212.09379)].
- [18] ATLAS collaboration, *The ATLAS collaboration software and firmware*, Tech. Rep. ATL-SOFT-PUB-2021-001, CERN, Geneva, Switzerland (2021), <https://cds.cern.ch/record/2767187>.
- [19] ATLAS collaboration, *Multijet simulation for 13 TeV ATLAS analyses*, Tech. Rep. ATL-PHYS-PUB-2019-017, CERN, Geneva, Switzerland (2019), <https://cds.cern.ch/record/2672252>.
- [20] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015), doi:[10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024) [preprint doi:[10.48550/arXiv.1410.3012](https://doi.org/10.48550/arXiv.1410.3012)].
- [21] ATLAS collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Switzerland (2014), <https://cds.cern.ch/record/1966419>.
- [22] NNPDF collaboration: R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867**, 244 (2013), doi:[10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003) [preprint doi:[10.48550/arXiv.1207.1303](https://doi.org/10.48550/arXiv.1207.1303)].
- [23] S. Agostinelli et al., *Geant4 — A simulation toolkit*, Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. **506**, 250 (2003), doi:[10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [24] ATLAS collaboration, *The ATLAS simulation infrastructure*, Eur. Phys. J. C **70**, 823 (2010), doi:[10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9) [preprint doi:[10.48550/arXiv.1005.4568](https://doi.org/10.48550/arXiv.1005.4568)].
- [25] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178**, 852 (2008), doi:[10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036) [preprint doi:[10.48550/arXiv.0710.3820](https://doi.org/10.48550/arXiv.0710.3820)].
- [26] ATLAS collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, Tech. Rep. ATL-PHYS-PUB-2016-017, CERN, Geneva, Switzerland (2016), <https://cds.cern.ch/record/2206965>.
- [27] ATLAS collaboration, *Vertex reconstruction performance of the ATLAS detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2015-026, CERN, Geneva, Switzerland (2015), <https://cds.cern.ch/record/2037717>.
- [28] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, J. High Energy Phys. **04**, 063 (2008), doi:[10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063) [preprint doi:[10.48550/arXiv.0802.1189](https://doi.org/10.48550/arXiv.0802.1189)].
- [29] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72**, 1896 (2012), doi:[10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2) [preprint doi:[10.48550/arXiv.1111.6097](https://doi.org/10.48550/arXiv.1111.6097)].
- [30] CMS collaboration, *Performance of quark/gluon discrimination using pp collisions at $\sqrt{s} = 8$ TeV*, Tech. Rep. CMS-PAS-JME-13-002, CERN, Geneva, Switzerland (2013), <https://cds.cern.ch/record/1599732>.

- [31] ATLAS collaboration, *Operation and performance of the ATLAS tile calorimeter in LHC Run 2*, Eur. Phys. J. C **84**, 1313 (2024), doi:[10.1140/epjc/s10052-024-13151-4](https://doi.org/10.1140/epjc/s10052-024-13151-4) [preprint doi:[10.48550/arXiv.2401.16034](https://doi.org/10.48550/arXiv.2401.16034)].
- [32] L. D. Landau, *On the energy loss of fast particles by ionisation*, in *Collected papers of L. D. Landau*, Elsevier, Amsterdam, Netherlands, ISBN 9780080105864 (1965), doi:[10.1016/b978-0-08-010586-4.50061-4](https://doi.org/10.1016/b978-0-08-010586-4.50061-4).
- [33] P. V. Vasilov, *Ionization losses of high-energy heavy particles*, Sov. J. Exp. Theor. Phys. **5**, 749 (1957).
- [34] A. Paszke et al., *PyTorch: An imperative style, high-performance deep learning library*, (arXiv preprint) doi:[10.48550/arXiv.1912.01703](https://doi.org/10.48550/arXiv.1912.01703).
- [35] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, (arXiv preprint) doi:[10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980).
- [36] J. S. Bridle, *Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters*, Adv. Neural Inf. Process. Syst. **2**, 211 (1989), https://proceedings.neurips.cc/paper_files/paper/1989/file/0336dcbab05b9d5ad24f4333c7658a0e-Paper.pdf.
- [37] J. S. Bridle, *Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition*, in *Neurocomputing*, Springer, Berlin, Heidelberg, Germany, ISBN 9783642761553 (1990), doi:[10.1007/978-3-642-76153-9_28](https://doi.org/10.1007/978-3-642-76153-9_28).
- [38] S. Kullback and R. A. Leibler, *On information and sufficiency*, Ann. Math. Stat. **22**, 79 (1951), doi:[10.1214/aoms/1177729694](https://doi.org/10.1214/aoms/1177729694).
- [39] T. Hastie, *Ridge regularization: An essential concept in data science*, Technometrics **62**, 426 (2020), doi:[10.1080/00401706.2020.1791959](https://doi.org/10.1080/00401706.2020.1791959).
- [40] P. Adragna et al., *Testbeam studies of production modules of the ATLAS Tile calorimeter*, Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. **606**, 362 (2009), doi:[10.1016/j.nima.2009.04.009](https://doi.org/10.1016/j.nima.2009.04.009).
- [41] Z. Ajaltouni et al., *Response of the ATLAS tile calorimeter prototype to muons*, Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. **388**, 64 (1997), doi:[10.1016/S0168-9002\(97\)00331-8](https://doi.org/10.1016/S0168-9002(97)00331-8).
- [42] R. Wigmans, *On the energy resolution of uranium and other hadron calorimeters*, Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. **259**, 389 (1987), doi:[10.1016/0168-9002\(87\)90823-0](https://doi.org/10.1016/0168-9002(87)90823-0).
- [43] Y. Gal and Z. Ghahramani, *Dropout as a Bayesian approximation: Representing model uncertainty in deep learning*, (arXiv preprint) doi:[10.48550/arXiv.1506.02142](https://doi.org/10.48550/arXiv.1506.02142).
- [44] M. Abdar et al., *A review of uncertainty quantification in deep learning: Techniques, applications and challenges*, Inf. Fusion **76**, 243 (2021), doi:[10.1016/j.inffus.2021.05.008](https://doi.org/10.1016/j.inffus.2021.05.008) [preprint doi:[10.48550/arXiv.2011.06225](https://doi.org/10.48550/arXiv.2011.06225)].
- [45] ATLAS collaboration, *ATLAS computing acknowledgements*, Tech. Rep. ATL-SOFT-PUB-2023-001, CERN, Geneva, Switzerland (2023), <https://cds.cern.ch/record/2869272>.

- [46] G. E. Hinton and D. van Camp, *Keeping the neural networks simple by minimizing the description length of the weights*, in *Proceedings of the sixth annual conference on computational learning theory*, Association for Computing Machinery, New York, USA, ISBN 9780897916110 (1993), doi:[10.1145/168304.168306](https://doi.org/10.1145/168304.168306).
- [47] D. M. Blei, A. Kucukelbir and J. D. McAuliffe, *Variational inference: A review for statisticians*, *J. Am. Stat. Assoc.* **112**, 859 (2017), doi:[10.1080/01621459.2017.1285773](https://doi.org/10.1080/01621459.2017.1285773) [preprint doi:[10.48550/arXiv.1601.00670](https://doi.org/10.48550/arXiv.1601.00670)].
- [48] L. P. Kadanoff, *Statistical physics*, World Scientific, Singapore, ISBN 9789810237585 (2000), doi:[10.1142/4016](https://doi.org/10.1142/4016).
- [49] Q. Liu and D. Wang, *Stein variational gradient descent: A general purpose Bayesian inference algorithm*, (arXiv preprint) doi:[10.48550/arXiv.1608.04471](https://doi.org/10.48550/arXiv.1608.04471).

The ATLAS collaboration

G. Aad¹⁰⁴, E. Aakvaag¹⁷, B. Abbott¹²³, S. Abdelhameed^{119a}, K. Abeling⁵⁶,
 N.J. Abicht⁵⁰, S.H. Abidi³⁰, M. Aboeela⁴⁶, A. Aboulhorma^{36e}, H. Abramowicz¹⁵⁵,
 Y. Abulaiti¹²⁰, B.S. Acharya^{70a,70b,n}, A. Ackermann^{64a}, C. Adam Bourdarios⁴,
 L. Adamczyk^{87a}, S.V. Addepalli¹⁴⁷, M.J. Addison¹⁰³, J. Adelman¹¹⁸, A. Adiguzel^{22c},
 T. Adye¹³⁷, A.A. Affolder¹³⁹, Y. Afik⁴¹, M.N. Agaras¹³, A. Aggarwal¹⁰²,
 C. Agheorghiesei^{28c}, F. Ahmadov^{40,ad}, S. Ahuja⁹⁷, X. Ai^{63e}, G. Aielli^{77a,77b},
 A. Aikot¹⁶⁸, M. Ait Tamlihat^{36e}, B. Aitbenkikh^{36a}, M. Akbiyik¹⁰², T.P.A. Åkesson¹⁰⁰,
 A.V. Akimov¹⁴⁹, D. Akiyama¹⁷³, N.N. Akolkar²⁵, S. Aktas^{22a}, G.L. Alberghi^{24b},
 J. Albert¹⁷⁰, P. Albicocco⁵⁴, G.L. Albouy⁶¹, S. Alderweireldt⁵³, Z.L. Alegria¹²⁴,
 M. Aleksa³⁷, I.N. Aleksandrov⁴⁰, C. Alexa^{28b}, T. Alexopoulos¹⁰, F. Alfonsi^{24b},
 M. Algren⁵⁷, M. Alhroob¹⁷², B. Ali¹³⁵, H.M.J. Ali^{93,w}, S. Ali³², S.W. Alibocus⁹⁴,
 M. Aliev^{34c}, G. Alimonti^{72a}, W. Alkakh⁵⁶, C. Allaire⁶⁷, B.M.M. Allbrooke¹⁵⁰,
 J.S. Allen¹⁰³, J.F. Allen⁵³, P.P. Allport²¹, A. Aloisio^{73a,73b}, F. Alonso⁹²,
 C. Alpigiani¹⁴², Z.M.K. Alsolami⁹³, A. Alvarez Fernandez¹⁰², M. Alves Cardoso⁵⁷,
 M.G. Alviggi^{73a,73b}, M. Aly¹⁰³, Y. Amaral Coutinho^{84b}, A. Ambler¹⁰⁶, C. Amelung³⁷,
 M. Ameri¹⁰³, C.G. Ames¹¹¹, D. Amidei¹⁰⁸, B. Amini⁵⁵, K.J. Amirie¹⁵⁹,
 A. Amirkhanov⁴⁰, S.P. Amor Dos Santos^{133a}, K.R. Amos¹⁶⁸, D. Amperiadou¹⁵⁶,
 S. An⁸⁵, V. Ananiev¹²⁸, C. Anastopoulos¹⁴³, T. Andeen¹¹, J.K. Anders⁹⁴,
 A.C. Anderson⁶⁰, A. Andreazza^{72a,72b}, S. Angelidakis⁹, A. Angerami⁴³,
 A.V. Anisenkov⁴⁰, A. Annovi^{75a}, C. Antel⁵⁷, E. Antipov¹⁴⁹, M. Antonelli⁵⁴,
 F. Anulli^{76a}, M. Aoki⁸⁵, T. Aoki¹⁵⁷, M.A. Aparo¹⁵⁰, L. Aperio Bella⁴⁹, C. Appelt¹⁵⁵,
 A. Apyan²⁷, S.J. Arbiol Val⁸⁸, C. Arcangeletti⁵⁴, A.T.H. Arce⁵², J-F. Arguin¹¹⁰,
 S. Argyropoulos¹⁵⁶, J.-H. Arling⁴⁹, O. Arnaez⁴, H. Arnold¹⁴⁹, G. Artoni^{76a,76b},
 H. Asada¹¹³, K. Asai¹²¹, S. Asai¹⁵⁷, N.A. Asbah³⁷, R.A. Ashby Pickering¹⁷²,
 A.M. Aslam⁹⁷, K. Assamagan³⁰, R. Astalos^{29a}, K.S.V. Astrand¹⁰⁰, S. Atashi¹⁶³,
 R.J. Atkin^{34a}, H. Atmani^{36f}, P.A. Atmasiddha¹³¹, K. Augsten¹³⁵, A.D. Aurioi⁴²,
 V.A. Austrup¹⁰³, G. Avolio³⁷, K. Axiotis⁵⁷, G. Azuelos^{110,ah}, D. Babal^{29b},
 H. Bachacou¹³⁸, K. Bachas^{156,r}, A. Bachi³⁵, E. Bachmann⁵¹, M.J. Backes^{64a},
 A. Badea⁴¹, T.M. Baer¹⁰⁸, P. Bagnaia^{76a,76b}, M. Bahmani¹⁹, D. Bahner⁵⁵, K. Bai¹²⁶,
 J.T. Baines¹³⁷, L. Baines⁹⁶, O.K. Baker¹⁷⁷, E. Bakos¹⁶, D. Bakshi Gupta⁸,
 L.E. Balabram Filho^{84b}, V. Balakrishnan¹²³, R. Balasubramanian⁴, E.M. Baldwin³⁹,
 P. Balek^{87a}, E. Ballabene^{24b,24a}, F. Balli¹³⁸, L.M. Baites^{64a}, W.K. Balunas³³,
 J. Balz¹⁰², I. Bamwidhi^{119b}, E. Banas⁸⁸, M. Bandieramonte¹³²,
 A. Bandyopadhyay²⁵, S. Bansal²⁵, L. Barak¹⁵⁵, M. Barakat⁴⁹, E.L. Barberio¹⁰⁷,
 D. Barberis^{58b,58a}, M. Barbero¹⁰⁴, M.Z. Barel¹¹⁷, T. Barillari¹¹², M-S. Barisits³⁷,
 T. Barklow¹⁴⁷, P. Baron¹²⁵, D.A. Baron Moreno¹⁰³, A. Baroncelli^{63a}, A.J. Barr¹²⁹,
 J.D. Barr⁹⁸, F. Barreiro¹⁰¹, J. Barreiro Guimarães da Costa¹⁴, M.G. Barros Teixeira^{133a},
 S. Barsov³⁹, F. Bartels^{64a}, R. Bartoldus¹⁴⁷, A.E. Barton⁹³, P. Bartos^{29a}, A. Basan¹⁰²,
 M. Baselga⁵⁰, S. Bashiri⁸⁸, A. Bassalat^{67,b}, M.J. Basso^{160a}, S. Bataju⁴⁶, R. Bate¹⁶⁹,
 R.L. Bates⁶⁰, S. Batlamous¹⁰¹, M. Battaglia¹³⁹, D. Battulga¹⁹, M. Bauge^{76a,76b},
 M. Bauer⁸⁰, P. Bauer²⁵, L.T. Bayer⁴⁹, L.T. Bazzano Hurrell³¹, J.B. Beacham¹¹²,
 T. Beau¹³⁰, J.Y. Beaucamp⁹², P.H. Beauchemin¹⁶², P. Bechtel²⁵, H.P. Beck^{20,q},
 K. Becker¹⁷², A.J. Beddall⁸³, V.A. Bednyakov⁴⁰, C.P. Bee¹⁴⁹, L.J. Beemster¹⁶,
 M. Begalli^{84d}, M. Begel³⁰, J.K. Behr⁴⁹, J.F. Beirer³⁷, F. Beisiegel²⁵, M. Belfkir^{119b},
 G. Bella¹⁵⁵, L. Bellagamba^{24b}, A. Bellerive³⁵, P. Bellos²¹, K. Beloborodov³⁹,
 D. Bencheikroun^{36a}, F. Bendebba^{36a}, Y. Benhammou¹⁵⁵, K.C. Benkendorfer⁶²,
 L. Beresford⁴⁹, M. Beretta⁵⁴, E. Bergeas Kuutmann¹⁶⁶, N. Berger⁴,
 B. Bergmann¹³⁵, J. Beringer^{18a}, G. Bernardi⁵, C. Bernius¹⁴⁷, F.U. Bernlochner²⁵,

- F. Bernon³⁷, A. Berrocal Guardia¹³, T. Berry⁹⁷, P. Berta¹³⁶, A. Berthold⁵¹,
 S. Bethke¹¹², A. Betti^{76a,76b}, A.J. Bevan⁹⁶, N.K. Bhalla⁵⁵, S. Bharthuar¹¹²,
 S. Bhatta¹⁴⁹, D.S. Bhattacharya¹⁷¹, P. Bhattacharai¹⁴⁷, Z.M. Bhatti¹²⁰, K.D. Bhide⁵⁵,
 V.S. Bhopatkar¹²⁴, R.M. Bianchi¹³², G. Bianco^{24b,24a}, O. Biebel¹¹¹, M. Biglietti^{78a},
 C.S. Billingsley⁴⁶, Y. Bimgdi^{36f}, M. Bindi⁵⁶, A. Bingham¹⁷⁶, A. Bingul^{22b},
 C. Bini^{76a,76b}, G.A. Bird³³, M. Birman¹⁷⁴, M. Biros¹³⁶, S. Biryukov¹⁵⁰, T. Bisanz⁵⁰,
 E. Bisceglie^{45b,45a}, J.P. Biswal¹³⁷, D. Biswas¹⁴⁵, I. Bloch⁴⁹, A. Blue⁶⁰,
 U. Blumenschein⁹⁶, J. Blumenthal¹⁰², V.S. Bobrovnikov⁴⁰, M. Boehler⁵⁵,
 B. Boehm¹⁷¹, D. Bogavac³⁷, A.G. Bogdanchikov³⁹, L.S. Boggia¹³⁰, V. Boisvert⁹⁷,
 P. Bokan³⁷, T. Bold^{87a}, M. Bomben⁵, M. Bona⁹⁶, M. Boonekamp¹³⁸,
 A.G. Borbély⁶⁰, I.S. Bordulev³⁹, G. Borissov⁹³, D. Bortoletto¹²⁹, D. Boscherini^{24b},
 M. Bosman¹³, K. Bouaouda^{36a}, N. Bouchhar¹⁶⁸, L. Boudet⁴, J. Boudreau¹³²,
 E.V. Bouhova-Thacker⁹³, D. Boumediene⁴², R. Bouquet^{58b,58a}, A. Boveia¹²²,
 J. Boyd³⁷, D. Boye³⁰, I.R. Boyko⁴⁰, L. Bozianu⁵⁷, J. Bracinik²¹, N. Brahimi⁴,
 G. Brandt¹⁷⁶, O. Brandt³³, B. Brau¹⁰⁵, J.E. Brau¹²⁶, R. Brenner¹⁷⁴, L. Brenner¹¹⁷,
 R. Brenner¹⁶⁶, S. Bressler¹⁷⁴, G. Brianti^{79a,79b}, D. Britton⁶⁰, D. Britzger¹¹²,
 I. Brock²⁵, R. Brock¹⁰⁹, G. Brooijmans⁴³, A.J. Brooks⁶⁹, E.M. Brooks^{160b}, E. Brost³⁰,
 L.M. Brown¹⁷⁰, L.E. Bruce⁶², T.L. Bruckler¹²⁹, P.A. Bruckman de Renstrom⁸⁸,
 B. Brüers⁴⁹, A. Bruni^{24b}, G. Bruni^{24b}, D. Brunner^{48a,48b}, M. Bruschi^{24b},
 N. Bruscinò^{76a,76b}, T. Buanes¹⁷, Q. Buat¹⁴², D. Buchin¹¹², A.G. Buckley⁶⁰,
 O. Bulekov³⁹, B.A. Bullard¹⁴⁷, S. Burdin⁹⁴, C.D. Burgard⁵⁰, A.M. Burger³⁷,
 B. Burghgrave⁸, O. Burlayenko⁵⁵, J. Burleson¹⁶⁷, J.T.P. Burr³³, J.C. Burzynski¹⁴⁶,
 E.L. Busch⁴³, V. Büscher¹⁰², P.J. Bussey⁶⁰, J.M. Butler²⁶, C.M. Buttar⁶⁰,
 J.M. Butterworth⁹⁸, W. Buttinger¹³⁷, C.J. Buxo Vazquez¹⁰⁹, A.R. Buzykaev⁴⁰,
 S. Cabrera Urbán¹⁶⁸, L. Cadamuro⁶⁷, D. Caforio⁵⁹, H. Cai¹³², Y. Cai^{24b,114c,24a},
 Y. Cai^{114a}, V.M.M. Cairo³⁷, O. Cakir^{3a}, N. Calace³⁷, P. Calafiura^{18a}, G. Calderini¹³⁰,
 P. Calfayan³⁵, G. Callea⁶⁰, L.P. Caloba^{84b}, D. Calvet⁴², S. Calvet⁴²,
 R. Camacho Toro¹³⁰, S. Camarda³⁷, D. Camarero Munoz²⁷, P. Camarri^{77a,77b},
 M.T. Camerlingo^{73a,73b}, D. Cameron³⁷, C. Camincher¹⁷⁰, M. Campanelli⁹⁸,
 A. Camplani⁴⁴, V. Canale^{73a,73b}, A.C. Canbay^{3a}, E. Canonero⁹⁷, J. Cantero¹⁶⁸,
 Y. Cao¹⁶⁷, F. Capocasa²⁷, M. Capua^{45b,45a}, A. Carbone^{72a,72b}, R. Cardarelli^{77a},
 J.C.J. Cardenas⁸, M.P. Cardiff²⁷, G. Carducci^{45b,45a}, T. Carli³⁷, G. Carlino^{73a},
 J.I. Carlotto¹³, B.T. Carlson^{132,s}, E.M. Carlson¹⁷⁰, J. Carmignani⁹⁴,
 L. Carminati^{72a,72b}, A. Carnelli¹³⁸, M. Carnesale³⁷, S. Caron¹¹⁶, E. Carquin^{140f},
 I.B. Carr¹⁰⁷, S. Carrá^{72a}, G. Carratta^{24b,24a}, A.M. Carroll¹²⁶, M.P. Casado^{13,i},
 M. Caspar⁴⁹, F.L. Castillo⁴, L. Castillo Garcia¹³, V. Castillo Gimenez¹⁶⁸,
 N.F. Castro^{133a,133e}, A. Catinaccio³⁷, J.R. Catmore¹²⁸, T. Cavaliere⁴, V. Cavaliere³⁰,
 L.J. Caviedes Betancourt^{23b}, Y.C. Cekmecelioglu⁴⁹, E. Celebi⁸³, S. Cella³⁷,
 V. Cepaitis⁵⁷, K. Cerny¹²⁵, A.S. Cerqueira^{84a}, A. Cerri^{75a,75b}, L. Cerrito^{77a,77b},
 F. Cerutti^{18a}, B. Cervato¹⁴⁵, A. Cervelli^{24b}, G. Cesarini⁵⁴, S.A. Cetin⁸³,
 P.M. Chabrilat¹³⁰, J. Chan^{18a}, W.Y. Chan¹⁵⁷, J.D. Chapman³³, E. Chapon¹³⁸,
 B. Chargeishvili^{153b}, D.G. Charlton²¹, C. Chauhan¹³⁶, Y. Che^{114a}, S. Chekanov⁶,
 S.V. Chekulaev^{160a}, G.A. Chelkov^{40,a}, B. Chen¹⁵⁵, B. Chen¹⁷⁰, H. Chen^{114a},
 H. Chen³⁰, J. Chen^{63c}, J. Chen¹⁴⁶, M. Chen¹²⁹, S. Chen⁸⁹, S.J. Chen^{114a},
 X. Chen^{63c}, X. Chen^{15,ag}, C.L. Cheng¹⁷⁵, H.C. Cheng^{65a}, S. Cheong¹⁴⁷,
 A. Cheplakov⁴⁰, E. Cheremushkina⁴⁹, E. Cherepanova¹¹⁷, R. Cherkaoui El Moursli^{36e},
 E. Cheu⁷, K. Cheung⁶⁶, L. Chevalier¹³⁸, V. Chiarella⁵⁴, G. Chiarelli^{75a},
 N. Chiedde¹⁰⁴, G. Chiodini^{71a}, A.S. Chisholm²¹, A. Chitan^{28b}, M. Chitishvili¹⁶⁸,
 M.V. Chizhov^{40,t}, K. Choi¹¹, Y. Chou¹⁴², E.Y.S. Chow¹¹⁶, K.L. Chu¹⁷⁴, M.C. Chu^{65a},
 X. Chu^{14,114c}, Z. Chubinidze⁵⁴, J. Chudoba¹³⁴, J.J. Chwastowski⁸⁸, D. Cieri¹¹²,

- K.M. Ciesla^{87a}, V. Cindro⁹⁵, A. Ciocio^{18a}, F. Ciroto^{73a,73b}, Z.H. Citron¹⁷⁴,
 M. Citterio^{72a}, D.A. Ciubotaru^{28b}, A. Clark⁵⁷, P.J. Clark⁵³, N. Clarke Hall⁹⁸,
 C. Clarry¹⁵⁹, S.E. Clawson⁴⁹, C. Clement^{48a,48b}, Y. Coadou¹⁰⁴, M. Cobal^{70a,70c},
 A. Coccaro^{58b}, R.F. Coelho Barrue^{133a}, R. Coelho Lopes De Sa¹⁰⁵, S. Coelli^{72a},
 L.S. Colangeli¹⁵⁹, B. Cole⁴³, P. Collado Soto¹⁰¹, J. Collot⁶¹, P. Conde Muiño^{133a,133g},
 M.P. Connell^{34c}, S.H. Connell^{34c}, E.I. Conroy¹²⁹, F. Conventi^{73a,ai}, H.G. Cooke²¹,
 A.M. Cooper-Sarkar¹²⁹, F.A. Corchia^{24b,24a}, A. Cordeiro Oudot Choi¹³⁰, L.D. Corpe⁴²,
 M. Corradi^{76a,76b}, F. Corriveau^{106,ac}, A. Cortes-Gonzalez¹⁹, M.J. Costa¹⁶⁸,
 F. Costanza⁴, D. Costanzo¹⁴³, B.M. Cote¹²², J. Couthures⁴, G. Cowan⁹⁷,
 K. Cranmer¹⁷⁵, L. Cremer⁵⁰, D. Cremonini^{24b,24a}, S. Crépe-Renaudin⁶¹,
 F. Crescioli¹³⁰, M. Cristinziani¹⁴⁵, M. Cristoforetti^{79a,79b}, V. Croft¹¹⁷, J.E. Crosby¹²⁴,
 G. Crosetti^{45b,45a}, A. Cueto¹⁰¹, H. Cui⁹⁸, Z. Cui⁷, W.R. Cunningham⁶⁰,
 F. Curcio¹⁶⁸, J.R. Curran⁵³, P. Czodrowski³⁷, M.J. Da Cunha Sargedas De Sousa^{58b,58a},
 J.V. Da Fonseca Pinto^{84b}, C. Da Via¹⁰³, W. Dabrowski^{87a}, T. Dado³⁷, S. Dahbi¹⁵²,
 T. Dai¹⁰⁸, D. Dal Santo²⁰, C. Dallapiccola¹⁰⁵, M. Dam⁴⁴, G. D'amen³⁰,
 V. D'Amico¹¹¹, J. Damp¹⁰², J.R. Dandoy³⁵, D. Dannheim³⁷, M. Danninger¹⁴⁶,
 V. Dao¹⁴⁹, G. Darbo^{58b}, S.J. Das³⁰, F. Dattola⁴⁹, S. D'Auria^{72a,72b},
 A. D'Avanzo^{73a,73b}, T. Davidek¹³⁶, I. Dawson⁹⁶, H.A. Day-hall¹³⁵, K. De⁸,
 C. De Almeida Rossi¹⁵⁹, R. De Asmundis^{73a}, N. De Biase⁴⁹, S. De Castro^{24b,24a},
 N. De Groot¹¹⁶, P. de Jong¹¹⁷, H. De la Torre¹¹⁸, A. De Maria^{114a}, A. De Salvo^{76a},
 U. De Sanctis^{77a,77b}, F. De Santis^{71a,71b}, A. De Santo¹⁵⁰, J.B. De Vivie De Regie⁶¹,
 J. Debevc⁹⁵, D.V. Dedovich⁴⁰, J. Degens⁹⁴, A.M. Deiana⁴⁶, J. Del Peso¹⁰¹,
 L. Delagrangé¹³⁰, F. Deliot¹³⁸, C.M. Delitzsch⁵⁰, M. Della Pietra^{73a,73b},
 D. Della Volpe⁵⁷, A. Dell'Acqua³⁷, L. Dell'Asta^{72a,72b}, M. Delmastro⁴,
 C.C. Delogu¹⁰², P.A. Delsart⁶¹, S. Demers¹⁷⁷, M. Demichev⁴⁰, S.P. Denisov³⁹,
 H. Denizli^{22a,1}, L. D'Eramo⁴², D. Derendarz⁸⁸, F. Derue¹³⁰, P. Dervan⁹⁴, K. Desch²⁵,
 C. Deutsch²⁵, F.A. Di Bello^{58b,58a}, A. Di Ciaccio^{77a,77b}, L. Di Ciaccio⁴,
 A. Di Domenico^{76a,76b}, C. Di Donato^{73a,73b}, A. Di Girolamo³⁷, G. Di Gregorio³⁷,
 A. Di Luca^{79a,79b}, B. Di Micco^{78a,78b}, R. Di Nardo^{78a,78b}, K.F. Di Petrillo⁴¹,
 M. Diamantopoulou³⁵, F.A. Dias¹¹⁷, T. Dias Do Vale¹⁴⁶, M.A. Diaz^{140a,140b},
 A.R. Didenko⁴⁰, M. Didenko¹⁶⁸, E.B. Diehl¹⁰⁸, S. Díez Cornell⁴⁹, C. Díez Pardo¹⁴⁵,
 C. Dimitriadi¹⁴⁸, A. Dimitrievska²¹, A. Dimri¹⁴⁹, J. Dingfelder²⁵, T. Dingley¹²⁹,
 I.-M. Dinu^{28b}, S.J. Dittmeier^{64b}, F. Dittus³⁷, M. Divisek¹³⁶, B. Dixit⁹⁴, F. Djama¹⁰⁴,
 T. Djobava^{153b}, C. Doglioni^{103,100}, A. Dohnalova^{29a}, Z. Dolezal¹³⁶, K. Domijan^{87a},
 K.M. Dona⁴¹, M. Donadelli^{84d}, B. Dong¹⁰⁹, J. Donini⁴², A. D'Onofrio^{73a,73b},
 M. D'Onofrio⁹⁴, J. Dopke¹³⁷, A. Doria^{73a}, N. Dos Santos Fernandes^{133a},
 P. Dougan¹⁰³, M.T. Dova⁹², A.T. Doyle⁶⁰, M.A. Draguet¹²⁹, M.P. Drescher⁵⁶,
 E. Dreyer¹⁷⁴, I. Drivas-koulouris¹⁰, M. Drnevic¹²⁰, M. Drozdova⁵⁷, D. Du^{63a},
 T.A. du Pree¹¹⁷, F. Dubinin³⁹, M. Dubovsky^{29a}, E. Duchovni¹⁷⁴, G. Duckeck¹¹¹,
 P.K. Duckett⁹⁸, O.A. Ducu^{28b}, D. Duda⁵³, A. Dudarev³⁷, E.R. Duden²⁷,
 M. D'uffizi¹⁰³, L. Duflot⁶⁷, M. Dührssen³⁷, I. Duminica^{28g}, A.E. Dumitriu^{28b},
 M. Dunford^{64a}, S. Dungs⁵⁰, K. Dunne^{48a,48b}, A. Duperrin¹⁰⁴, H. Duran Yildiz^{3a},
 M. Düren⁵⁹, A. Durglishvili^{153b}, D. Duvnjak³⁵, B.L. Dwyer¹¹⁸, G.I. Dyckes^{18a},
 M. Dyndal^{87a}, B.S. Dziedzic³⁷, Z.O. Earnshaw¹⁵⁰, G.H. Eberwein¹²⁹, B. Eckerova^{29a},
 S. Eggebrecht⁵⁶, E. Egidio Purcino De Souza^{84e}, G. Eigen¹⁷, K. Einsweiler^{18a},
 T. Ekelof¹⁶⁶, P.A. Ekman¹⁰⁰, S. El Farkh^{36b}, Y. El Ghazali^{63a}, H. El Jarrari³⁷,
 A. El Moussaouy^{36a}, V. Ellajosyula¹⁶⁶, M. Ellert¹⁶⁶, F. Ellinghaus¹⁷⁶, N. Ellis³⁷,
 J. Elmsheuser³⁰, M. Elsayy^{119a}, M. Elsing³⁷, D. Emeliyanov¹³⁷, Y. Enari⁸⁵,
 I. Ene^{18a}, S. Epari¹³, D. Ernani Martins Neto⁸⁸, M. Errenst¹⁷⁶, M. Escalier⁶⁷,
 C. Escobar¹⁶⁸, E. Etzion¹⁵⁵, G. Evans^{133a,133b}, H. Evans⁶⁹, L.S. Evans⁹⁷,

- A. Ezhilov³⁹, S. Ezzarqtouni^{36a}, F. Fabbri^{24b,24a}, L. Fabbri^{24b,24a}, G. Facini⁹⁸,
 V. Fadeyev¹³⁹, R.M. Fakhruddinov³⁹, D. Fakoudis¹⁰², S. Falciano^{76a},
 L.F. Falda Ulhoa Coelho^{133a}, F. Fallavollita¹¹², G. Falsetti^{45b,45a}, J. Faltova¹³⁶,
 C. Fan¹⁶⁷, K.Y. Fan^{65b}, Y. Fan¹⁴, Y. Fang^{14,114c}, M. Fanti^{72a,72b}, M. Faraj^{70a,70b},
 Z. Farazpay⁹⁹, A. Farbin⁸, A. Farilla^{78a}, T. Farooque¹⁰⁹, J.N. Farr¹⁷⁷,
 S.M. Farrington^{137,53}, F. Fassi^{36e}, D. Fassouliotis⁹, L. Fayard⁶⁷, P. Federic¹³⁶,
 P. Federicova¹³⁴, O.L. Fedin^{39,a}, M. Feickert¹⁷⁵, L. Feligioni¹⁰⁴, D.E. Fellers^{18a},
 C. Feng^{63b}, Z. Feng¹¹⁷, M.J. Fenton¹⁶³, L. Ferencz⁴⁹, P. Fernandez Martinez⁶⁸,
 M.J.V. Fernoux¹⁰⁴, J. Ferrando⁹³, A. Ferrari¹⁶⁶, P. Ferrari^{117,116}, R. Ferrari^{74a},
 D. Ferrere⁵⁷, C. Ferretti¹⁰⁸, M.P. Fewell¹, D. Fiacco^{76a,76b}, F. Fiedler¹⁰²,
 P. Fiedler¹³⁵, S. Filimonov³⁹, A. Filipčić⁹⁵, E.K. Filmer^{160a}, F. Filthaut¹¹⁶,
 M.C.N. Fiolhais^{133a,133c,c}, L. Fiorini¹⁶⁸, W.C. Fisher¹⁰⁹, T. Fitschen¹⁰³,
 P.M. Fitzhugh¹³⁸, I. Fleck¹⁴⁵, P. Fleischmann¹⁰⁸, T. Flick¹⁷⁶, M. Flores^{34d,ae},
 L.R. Flores Castillo^{65a}, L. Flores Sanz De Acedo³⁷, F.M. Follega^{79a,79b}, N. Fomin³³,
 J.H. Foo¹⁵⁹, A. Formica¹³⁸, A.C. Forti¹⁰³, E. Fortin³⁷, A.W. Fortman^{18a},
 L. Fountas^{9,j}, D. Fournier⁶⁷, H. Fox⁹³, P. Francavilla^{75a,75b}, S. Francescato⁶²,
 S. Franchellucci⁵⁷, M. Franchini^{24b,24a}, S. Franchino^{64a}, D. Francis³⁷, L. Franco¹¹⁶,
 V. Franco Lima³⁷, L. Franconi⁴⁹, M. Franklin⁶², G. Frattari²⁷, Y.Y. Frid¹⁵⁵,
 J. Friend⁶⁰, N. Fritzsche³⁷, A. Froch⁵⁷, D. Froidevaux³⁷, J.A. Frost¹²⁹, Y. Fu¹⁰⁹,
 S. Fuenzalida Garrido^{140f}, M. Fujimoto¹⁰⁴, K.Y. Fung^{65a}, E. Furtado De Simas Filho^{84e},
 M. Furukawa¹⁵⁷, J. Fuster¹⁶⁸, A. Gaa⁵⁶, A. Gabrielli^{24b,24a}, A. Gabrielli¹⁵⁹,
 P. Gadow³⁷, G. Gagliardi^{58b,58a}, L.G. Gagnon^{18a}, S. Gaid¹⁶⁵, S. Galantzan¹⁵⁵,
 J. Gallagher¹, E.J. Gallas¹²⁹, A.L. Gallen¹⁶⁶, B.J. Gallop¹³⁷, K.K. Gan¹²²,
 S. Ganguly¹⁵⁷, Y. Gao⁵³, A. Garabaglu¹⁴², F.M. Garay Walls^{140a,140b}, B. Garcia³⁰,
 C. García¹⁶⁸, A. Garcia Alonso¹¹⁷, A.G. Garcia Caffaro¹⁷⁷, J.E. García Navarro¹⁶⁸,
 M. Garcia-Sciveres^{18a}, G.L. Gardner¹³¹, R.W. Gardner⁴¹, N. Garelli¹⁶², R.B. Garg¹⁴⁷,
 J.M. Gargan⁵³, C.A. Garner¹⁵⁹, C.M. Garvey^{34a}, V.K. Gassmann¹⁶², G. Gaudio^{74a},
 V. Gautam¹³, P. Gauzzi^{76a,76b}, J. Gavranovic⁹⁵, I.L. Gavrilenko³⁹, A. Gavriluk³⁹,
 C. Gay¹⁶⁹, G. Gaycken¹²⁶, E.N. Gazis¹⁰, A. Gekow¹²², C. Gemme^{58b}, M.H. Genest⁶¹,
 A.D. Gentry¹¹⁵, S. George⁹⁷, W.F. George²¹, T. Gerialis⁴⁷, A.A. Gerwin¹²³,
 P. Gessinger-Befurt³⁷, M.E. Geyik¹⁷⁶, M. Ghani¹⁷², K. Ghorbanian⁹⁶, A. Ghosal¹⁴⁵,
 A. Ghosh¹⁶³, A. Ghosh⁷, B. Giacobbe^{24b}, S. Giagu^{76a,76b}, T. Giani¹¹⁷,
 A. Giannini^{63a}, S.M. Gibson⁹⁷, M. Gignac¹³⁹, D.T. Gil^{87b}, A.K. Gilbert^{87a},
 B.J. Gilbert⁴³, D. Gillberg³⁵, G. Gilles¹¹⁷, L. Ginabat¹³⁰, D.M. Gingrich^{2,ah},
 M.P. Giordani^{70a,70c}, P.F. Giraud¹³⁸, G. Giugliarelli^{70a,70c}, D. Giugni^{72a}, F. Giuli^{77a,77b},
 I. Gkialas^{9,j}, L.K. Gladilin³⁹, C. Glasman¹⁰¹, G. Glemža⁴⁹, M. Glisic¹²⁶, I. Gnesi^{45b},
 Y. Go³⁰, M. Goblirsch-Kolb³⁷, B. Gocke⁵⁰, D. Godin¹¹⁰, B. Gokturk^{22a},
 S. Goldfarb¹⁰⁷, T. Golling⁵⁷, M.G.D. Gololo^{34c}, D. Golubkov³⁹, J.P. Gombas¹⁰⁹,
 A. Gomes^{133a,133b}, G. Gomes Da Silva¹⁴⁵, A.J. Gomez Delegido¹⁶⁸, R. Gonçalves^{133a},
 L. Gonella²¹, A. Gongadze^{153c}, F. Gonnella²¹, J.L. Gonski¹⁴⁷,
 R.Y. González Andana⁵³, S. González de la Hoz¹⁶⁸, R. Gonzalez Lopez⁹⁴,
 C. Gonzalez Renteria^{18a}, M.V. Gonzalez Rodrigues⁴⁹, R. Gonzalez Suarez¹⁶⁶,
 S. Gonzalez-Sevilla⁵⁷, L. Goossens³⁷, B. Gorini³⁷, E. Gorini^{71a,71b}, A. Gorišek⁹⁵,
 T.C. Gosart¹³¹, A.T. Goshaw⁵², M.I. Gostkin⁴⁰, S. Goswami¹²⁴, C.A. Gottardo³⁷,
 S.A. Gotz¹¹¹, M. Gouighri^{36b}, A.G. Goussiou¹⁴², N. Govender^{34c}, R.P. Grabarczyk¹²⁹,
 I. Grabowska-Bold^{87a}, K. Graham³⁵, E. Gramstad¹²⁸, S. Grancagnolo^{71a,71b},
 C.M. Grant^{1,138}, P.M. Gravila^{28f}, F.G. Gravili^{71a,71b}, H.M. Gray^{18a}, M. Greco¹¹²,
 M.J. Green¹, C. Grefe²⁵, A.S. Grefsrud¹⁷, I.M. Gregor⁴⁹, K.T. Greif¹⁶³,
 P. Grenier¹⁴⁷, S.G. Grewe¹¹², A.A. Grillo¹³⁹, K. Grimm³², S. Grinstein^{13,x},
 J.-F. Grivaz⁶⁷, E. Gross¹⁷⁴, J. Grosse-Knetter⁵⁶, L. Guan¹⁰⁸, G. Guerrieri³⁷,

R. Gugel¹⁰², J.A.M. Guhit¹⁰⁸, A. Guida¹⁹, E. Guilloton¹⁷², S. Guindon³⁷,
 F. Guo^{14,114c}, J. Guo^{63c}, L. Guo⁴⁹, L. Guo^{114b,v}, Y. Guo¹⁰⁸, A. Gupta⁵⁰,
 R. Gupta¹³², S. Gurbuz²⁵, S.S. Gurdasani⁴⁹, G. Gustavino^{76a,76b}, P. Gutierrez¹²³,
 L.F. Gutierrez Zagazeta¹³¹, M. Gutsche⁵¹, C. Gutschow⁹⁸, C. Gwenlan¹²⁹,
 C.B. Gwilliam⁹⁴, E.S. Haaland¹²⁸, A. Haas¹²⁰, M. Habedank⁶⁰, C. Haber^{18a},
 H.K. Hadavand⁸, A. Haddad⁴², A. Hadeef⁵¹, A.I. Hagan⁹³, J.J. Hahn¹⁴⁵,
 E.H. Haines⁹⁸, M. Haleem¹⁷¹, J. Haley¹²⁴, G.D. Hallewell¹⁰⁴, L. Halser²⁰,
 K. Hamano¹⁷⁰, M. Hamer²⁵, E.J. Hampshire⁹⁷, J. Han^{63b}, L. Han^{114a}, L. Han^{63a},
 S. Han^{18a}, K. Hanagaki⁸⁵, M. Hance¹³⁹, D.A. Hangal⁴³, H. Hanif¹⁴⁶,
 M.D. Hank¹³¹, J.B. Hansen⁴⁴, P.H. Hansen⁴⁴, D. Harada⁵⁷, T. Harenberg¹⁷⁶,
 S. Harkusha¹⁷⁸, M.L. Harris¹⁰⁵, Y.T. Harris²⁵, J. Harrison¹³, N.M. Harrison¹²²,
 P.F. Harrison¹⁷², N.M. Hartman¹¹², N.M. Hartmann¹¹¹, R.Z. Hasan^{97,137},
 Y. Hasegawa¹⁴⁴, F. Haslbeck¹²⁹, S. Hassan¹⁷, R. Hauser¹⁰⁹, C.M. Hawkes²¹,
 R.J. Hawkings³⁷, Y. Hayashi¹⁵⁷, D. Hayden¹⁰⁹, C. Hayes¹⁰⁸, R.L. Hayes¹¹⁷,
 C.P. Hays¹²⁹, J.M. Hays⁹⁶, H.S. Hayward⁹⁴, F. He^{63a}, M. He^{14,114c}, Y. He⁴⁹,
 Y. He⁹⁸, N.B. Heatley⁹⁶, V. Hedberg¹⁰⁰, A.L. Heggelund¹²⁸, C. Heidegger⁵⁵,
 K.K. Heidegger⁵⁵, J. Heilman³⁵, S. Heim⁴⁹, T. Heim^{18a}, T. Heimel^z, J.G. Heinlein¹³¹,
 J.J. Heinrich¹²⁶, L. Heinrich^{112,af}, J. Hejbal¹³⁴, A. Held¹⁷⁵, S. Hellesund¹⁷,
 C.M. Helling¹⁶⁹, S. Hellman^{48a,48b}, L. Henkelmann³³, A.M. Henriques Correia³⁷,
 H. Herde¹⁰⁰, Y. Hernández Jiménez¹⁴⁹, L.M. Herrmann²⁵, T. Herrmann⁵¹,
 G. Herten⁵⁵, R. Hertenberger¹¹¹, L. Hervas³⁷, M.E. Hespington¹⁰², N.P. Hessey^{160a},
 J. Hessler¹¹², M. Hidaoui^{36b}, N. Hidic¹³⁶, E. Hill¹⁵⁹, S.J. Hillier²¹, J.R. Hinds¹⁰⁹,
 F. Hinterkeuser²⁵, M. Hirose¹²⁷, S. Hirose¹⁶¹, D. Hirschbuehl¹⁷⁶, T.G. Hitchings¹⁰³,
 B. Hiti⁹⁵, J. Hobbs¹⁴⁹, R. Hobincu^{28e}, N. Hod¹⁷⁴, M.C. Hodgkinson¹⁴³,
 B.H. Hodgkinson¹²⁹, A. Hoecker³⁷, D.D. Hofer¹⁰⁸, J. Hofer¹⁶⁸, M. Holzbock³⁷,
 L.B.A.H. Hommels³³, B.P. Honan¹⁰³, J.J. Hong⁶⁹, J. Hong^{63c}, T.M. Hong¹³²,
 B.H. Hooberman¹⁶⁷, W.H. Hopkins⁶, M.C. Hoppesch¹⁶⁷, Y. Horii¹¹³,
 M.E. Horstmann¹¹², S. Hou¹⁵², M.R. Housenga¹⁶⁷, A.S. Howard⁹⁵, J. Howarth⁶⁰,
 J. Hoya⁶, M. Hrabovsky¹²⁵, T. Hryn'ova⁴, P.J. Hsu⁶⁶, S.-C. Hsu¹⁴², T. Hsu⁶⁷,
 M. Hu^{18a}, Q. Hu^{63a}, S. Huang³³, X. Huang^{14,114c}, Y. Huang¹³⁶, Y. Huang^{114b},
 Y. Huang¹⁰², Y. Huang¹⁴, Z. Huang¹⁰³, Z. Hubacek¹³⁵, M. Huebner²⁵,
 F. Huegging²⁵, T.B. Huffman¹²⁹, M. Hufnagel Maranha De Faria^{84a}, C.A. Hugli⁴⁹,
 M. Huhtinen³⁷, S.K. Huiberts¹⁷, R. Hulskens¹⁰⁶, C.E. Hultquist^{18a}, N. Huseynov^{12,g},
 J. Huston¹⁰⁹, J. Huth⁶², R. Hyneman⁷, G. Iacobucci⁵⁷, G. Iakovidis³⁰,
 L. Iconomidou-Fayard⁶⁷, J.P. Iddon³⁷, P. Iengo^{73a,73b}, R. Iguchi¹⁵⁷, Y. Iiyama¹⁵⁷,
 T. Iizawa¹²⁹, Y. Ikegami⁸⁵, D. Iliadis¹⁵⁶, N. Ilic¹⁵⁹, H. Imam^{84c},
 G. Inacio Goncalves^{84d}, S.A. Infante Cabanas^{140c}, T. Ingebretsen Carlson^{48a,48b},
 J.M. Inglis⁹⁶, G. Introzzi^{74a,74b}, M. Iodice^{78a}, V. Ippolito^{76a,76b}, R.K. Irwin⁹⁴,
 M. Ishino¹⁵⁷, W. Islam¹⁷⁵, C. Issever¹⁹, S. Istin^{22a,am}, H. Ito¹⁷³, R. Iuppa^{79a,79b},
 A. Ivina¹⁷⁴, V. Izzo^{73a}, P. Jacka¹³⁴, P. Jackson¹, P. Jain⁴⁹, K. Jakobs⁵⁵,
 T. Jakoubek¹⁷⁴, J. Jamieson⁶⁰, W. Jang¹⁵⁷, M. Javurkova¹⁰⁵, P. Jawahar¹⁰³,
 L. Jeanty¹²⁶, J. Jejelava^{153a}, P. Jenni^{55,f}, C.E. Jessiman³⁵, C. Jia^{63b}, H. Jia¹⁶⁹,
 J. Jia¹⁴⁹, X. Jia^{14,114c}, Z. Jia^{114a}, C. Jiang⁵³, Q. Jiang^{65b}, S. Jiggins⁴⁹,
 J. Jimenez Pena¹³, S. Jin^{114a}, A. Jinaru^{28b}, O. Jinnouchi¹⁴¹, P. Johansson¹⁴³,
 K.A. Johns⁷, J.W. Johnson¹³⁹, F.A. Jolly⁴⁹, D.M. Jones¹⁵⁰, E. Jones⁴⁹, K.S. Jones⁸,
 P. Jones³³, R.W.L. Jones⁹³, T.J. Jones⁹⁴, H.L. Joos^{56,37}, R. Joshi¹²², J. Jovicevic¹⁶,
 X. Ju^{18a}, J.J. Junggeburth³⁷, T. Junkermann^{64a}, A. Juste Rozas^{13,x}, M.K. Juzek⁸⁸,
 S. Kabana^{140e}, A. Kaczmarska⁸⁸, M. Kado¹¹², H. Kagan¹²², M. Kagan¹⁴⁷,
 A. Kahn¹³¹, C. Kahra¹⁰², T. Kaji¹⁵⁷, E. Kajomovitz¹⁵⁴, N. Kakati¹⁷⁴,
 I. Kalaitzidou⁵⁵, N.J. Kang¹³⁹, D. Kar^{34g}, K. Karava¹²⁹, E. Karentzos²⁵,

- O. Karkout¹¹⁷, S.N. Karpov⁴⁰, Z.M. Karpova⁴⁰, V. Kartvelishvili⁹³, A.N. Karyukhin³⁹,
 E. Kasimi¹⁵⁶, J. Katzy⁴⁹, S. Kaur³⁵, K. Kawade¹⁴⁴, M.P. Kawale¹²³, C. Kawamoto⁸⁹,
 T. Kawamoto^{63a}, E.F. Kay³⁷, F.I. Kaya¹⁶², S. Kazakos¹⁰⁹, V.F. Kazanin³⁹, Y. Ke¹⁴⁹,
 J.M. Keaveney^{34a}, R. Keeler¹⁷⁰, G.V. Kehris⁶², J.S. Keller³⁵, J.J. Kempster¹⁵⁰,
 O. Kepka¹³⁴, J. Kerri^{160b}, B.P. Kerridge¹³⁷, B.P. Kerševan⁹⁵, L. Keszeghova^{29a},
 R.A. Khan¹³², A. Khanov¹²⁴, A.G. Kharlamov³⁹, T. Kharlamova³⁹, E.E. Khoda¹⁴²,
 M. Kholodenko^{133a}, T.J. Khoo¹⁹, G. Khoriauli¹⁷¹, J. Khubua^{153b,*}, Y.A.R. Khwaira¹³⁰,
 B. Kibirige^{34g}, D. Kim⁶, D.W. Kim^{48a,48b}, Y.K. Kim⁴¹, N. Kimura⁹⁸, M.K. Kingston⁵⁶,
 A. Kirchoff⁵⁶, C. Kirfel²⁵, F. Kirfel²⁵, J. Kirk¹³⁷, A.E. Kiryunin¹¹², S. Kita¹⁶¹,
 C. Kitsaki¹⁰, O. Kivernyk²⁵, M. Klassen¹⁶², C. Klein³⁵, L. Klein¹⁷¹, M.H. Klein⁴⁶,
 S.B. Klein⁵⁷, U. Klein⁹⁴, A. Klimentov³⁰, T. Klioutchnikova³⁷, P. Kluit¹¹⁷,
 S. Kluth¹¹², E. Kneringer⁸⁰, T.M. Knight¹⁵⁹, A. Knue⁵⁰, D. Kobylanski¹⁷⁴,
 S.F. Koch¹²⁹, M. Kocian¹⁴⁷, P. Kodyš¹³⁶, D.M. Koeck¹²⁶, P.T. Koenig²⁵, T. Koffas³⁵,
 O. Kolay⁵¹, I. Koletsou⁴, T. Komarek⁸⁸, K. Köneke⁵⁶, A.X.Y. Kong¹, T. Kono¹²¹,
 N. Konstantinidis⁹⁸, P. Kontaxakis⁵⁷, B. Konya¹⁰⁰, R. Kopeliansky⁴³, S. Koperny^{87a},
 K. Korcyl⁸⁸, K. Kordas^{156,e}, A. Korn⁹⁸, S. Korn⁵⁶, I. Korolkov¹³, N. Korotkova³⁹,
 B. Kortman¹¹⁷, O. Kortner¹¹², S. Kortner¹¹², W.H. Kostecka¹¹⁸, V.V. Kostyukhin¹⁴⁵,
 A. Kotsokechagia³⁷, A. Kotwal⁵², A. Koulouris³⁷, A. Kourkoumeli-Charalampidi^{74a,74b},
 C. Kourkoumelis⁹, E. Kourlitis^{112,af}, O. Kovanda¹²⁶, R. Kowalewski¹⁷⁰,
 W. Kozanecki¹²⁶, A.S. Kozhin³⁹, V.A. Kramarenko³⁹, G. Kramberger⁹⁵, P. Kramer²⁵,
 M.W. Krasny¹³⁰, A. Krasznahorkay¹⁰⁵, A.C. Kraus¹¹⁸, J.W. Kraus¹⁷⁶, J.A. Kremer⁴⁹,
 T. Kresse⁵¹, L. Kretschmann¹⁷⁶, J. Kretzschmar⁹⁴, K. Kreul¹⁹, P. Krieger¹⁵⁹,
 K. Krizka²¹, K. Kroeninger⁵⁰, H. Kroha¹¹², J. Kroll¹³⁴, J. Kroll¹³¹,
 K.S. Krowpman¹⁰⁹, U. Kruchonak⁴⁰, H. Krüger²⁵, N. Krumnack⁸², M.C. Kruse⁵²,
 O. Kuchinskaia³⁹, S. Kuday^{3a}, S. Kuehn³⁷, R. Kuesters⁵⁵, T. Kuhl⁴⁹, V. Kukhtin⁴⁰,
 Y. Kulchitsky⁴⁰, S. Kuleshov^{140d,140b}, M. Kumar^{34g}, N. Kumari⁴⁹, P. Kumari^{160b},
 A. Kupco¹³⁴, T. Kupfer⁵⁰, A. Kupich³⁹, O. Kuprash⁵⁵, H. Kurashige⁸⁶,
 L.L. Kurchaninov^{160a}, O. Kurdysh⁴, Y.A. Kurochkin³⁸, A. Kurova³⁹, M. Kuze¹⁴¹,
 A.K. Kvam¹⁰⁵, J. Kvita¹²⁵, N.G. Kyriacou¹⁰⁸, L.A.O. Laatu¹⁰⁴, C. Lacasta¹⁶⁸,
 F. Lacava^{76a,76b}, H. Lacker¹⁹, D. Lacour¹³⁰, N.N. Lad⁹⁸, E. Ladygin⁴⁰, A. Lafarge⁴²,
 B. Laforge¹³⁰, T. Lagouri¹⁷⁷, F.Z. Lahbabi^{36a}, S. Lai⁵⁶, J.E. Lambert¹⁷⁰,
 S. Lammers⁶⁹, W. Lampl⁷, C. Lampoudis^{156,e}, G. Lamprinoudis¹⁰²,
 A.N. Lancaster¹¹⁸, E. Lançon³⁰, U. Landgraf⁵⁵, M.P.J. Landon⁹⁶, V.S. Lang⁵⁵,
 O.K.B. Langrekken¹²⁸, A.J. Lankford¹⁶³, F. Lanni³⁷, K. Lantzsch²⁵, A. Lanza^{74a},
 M. Lanzac Berrocal¹⁶⁸, J.F. Laporte¹³⁸, T. Lari^{72a}, F. Lasagni Manghi^{24b},
 M. Lassnig³⁷, V. Latonova¹³⁴, S.D. Lawlor¹⁴³, Z. Lawrence¹⁰³, R. Lazaridou¹⁷²,
 M. Lazzaroni^{72a,72b}, H.D.M. Le¹⁰⁹, E.M. Le Boulicaut¹⁷⁷, L.T. Le Pottier^{18a},
 B. Leban^{24b,24a}, M. LeBlanc¹⁰³, F. Ledroit-Guillon⁶¹, S.C. Lee¹⁵², T.F. Lee⁹⁴,
 L.L. Leeuw^{34c,ak}, M. Lefebvre¹⁷⁰, C. Leggett^{18a}, G. Lehmann Miotto³⁷, M. Leigh⁵⁷,
 W.A. Leight¹⁰⁵, W. Leinonen¹¹⁶, A. Leisos^{156,u}, M.A.L. Leite^{84c}, C.E. Leitgeb¹⁹,
 R. Leitner¹³⁶, K.J.C. Leney⁴⁶, T. Lenz²⁵, S. Leone^{75a}, C. Leonidopoulos⁵³,
 A. Leopold¹⁴⁸, J.H. Lepage Bourbonnais³⁵, R. Les¹⁰⁹, C.G. Lester³³,
 M. Levchenko³⁹, J. Levêque⁴, L.J. Levinson¹⁷⁴, G. Levrini^{24b,24a}, M.P. Lewicki⁸⁸,
 C. Lewis¹⁴², D.J. Lewis⁴, L. Lewitt¹⁴³, A. Li³⁰, B. Li^{63b}, C. Li¹⁰⁸, C-Q. Li¹¹²,
 H. Li^{63a}, H. Li^{63b}, H. Li¹⁰³, H. Li¹⁵, H. Li^{63b}, J. Li^{63c}, K. Li¹⁴, L. Li^{63c},
 R. Li¹⁷⁷, S. Li^{14,114c}, S. Li^{63d,63c,d}, T. Li⁵, X. Li¹⁰⁶, Z. Li¹⁵⁷, Z. Li^{14,114c}, Z. Li^{63a},
 S. Liang^{14,114c}, Z. Liang¹⁴, M. Liberatore¹³⁸, B. Liberti^{77a}, K. Lie^{65c},
 J. Lieber Marin^{84e}, H. Lien⁶⁹, H. Lin¹⁰⁸, L. Linden¹¹¹, R.E. Lindley⁷, J.H. Lindon²,
 J. Ling⁶², E. Lipeles¹³¹, A. Lipniacka¹⁷, A. Lister¹⁶⁹, J.D. Little⁶⁹, B. Liu¹⁴,
 B.X. Liu^{114b}, D. Liu^{63d,63c}, E.H.L. Liu²¹, J.K.K. Liu³³, K. Liu^{63d}, K. Liu^{63d,63c},

- M. Liu^{63a}, M.Y. Liu^{63a}, P. Liu¹⁴, Q. Liu^{63d,142,63c}, X. Liu^{63a}, X. Liu^{63b},
 Y. Liu^{114b,114c}, Y.L. Liu^{63b}, Y.W. Liu^{63a}, S.L. Lloyd⁹⁶, E.M. Lobodzinska⁴⁹, P. Loch⁷,
 E. Lodhi¹⁵⁹, T. Lohse¹⁹, K. Lohwasser¹⁴³, E. Loiacono⁴⁹, J.D. Lomas²¹,
 J.D. Long⁴³, I. Longarini¹⁶³, R. Longo¹⁶⁷, A. Lopez Solis⁴⁹, N.A. Lopez-canelas⁷,
 N. Lorenzo Martinez⁴, A.M. Lory¹¹¹, M. Losada^{119a}, G. Lösckce Centeno¹⁵⁰,
 O. Loseva³⁹, X. Lou^{48a,48b}, X. Lou^{14,114c}, A. Lounis⁶⁷, P.A. Love⁹³, G. Lu^{14,114c},
 M. Lu⁶⁷, S. Lu¹³¹, Y.J. Lu¹⁵², H.J. Lubatti¹⁴², C. Luci^{76a,76b}, F.L. Lucio Alves^{114a},
 F. Luehring⁶⁹, B.S. Lunday¹³¹, O. Lundberg¹⁴⁸, B. Lund-Jensen^{148,*}, N.A. Luongo⁶,
 M.S. Lutz³⁷, A.B. Lux²⁶, D. Lynn³⁰, R. Lysak¹³⁴, E. Lytken¹⁰⁰, V. Lyubushkin⁴⁰,
 T. Lyubushkina⁴⁰, M.M. Lyukova¹⁴⁹, M. Firdaus M. Soberi⁵³, H. Ma³⁰, K. Ma^{63a},
 L.L. Ma^{63b}, W. Ma^{63a}, Y. Ma¹²⁴, J.C. MacDonald¹⁰²,
 P.C. Machado De Abreu Farias^{84e}, R. Madar⁴², T. Madula⁹⁸, J. Maeda⁸⁶, T. Maeno³⁰,
 P.T. Mafa^{34c,k}, H. Maguire¹⁴³, V. Maiboroda¹³⁸, A. Maio^{133a,133b,133d}, K. Maj^{87a},
 O. Majersky⁴⁹, S. Majewski¹²⁶, R. Makhmanazarov³⁹, N. Makovec⁶⁷,
 V. Maksimovic¹⁶, B. Malaescu¹³⁰, Pa. Malecki⁸⁸, V.P. Maleev³⁹, F. Malek^{61,p},
 M. Mali⁹⁵, D. Malito⁹⁷, U. Mallik^{81,*}, S. Maltezos¹⁰, S. Malyukov⁴⁰, J. Mamuzic¹³,
 G. Mancini⁵⁴, M.N. Mancini²⁷, G. Manco^{74a,74b}, J.P. Mandalia⁹⁶, S.S. Mandarry¹⁵⁰,
 I. Mandić⁹⁵, L. Manhaes de Andrade Filho^{84a}, I.M. Maniatis¹⁷⁴, J. Manjarres Ramos⁹¹,
 D.C. Mankad¹⁷⁴, A. Mann¹¹¹, S. Manzoni³⁷, L. Mao^{63c}, X. Mapekula^{34c},
 A. Marantis^{156,u}, G. Marchiori⁵, M. Marcisovsky¹³⁴, C. Marcon^{72a}, M. Marinescu²¹,
 S. Marium⁴⁹, M. Marjanovic¹²³, A. Markhoos⁵⁵, M. Markovitch⁶⁷, M.K. Maroun¹⁰⁵,
 E.J. Marshall⁹³, Z. Marshall^{18a}, S. Marti-Garcia¹⁶⁸, J. Martin⁹⁸, T.A. Martin¹³⁷,
 V.J. Martin⁵³, B. Martin dit Latour¹⁷, L. Martinelli^{76a,76b}, M. Martinez^{13,x},
 P. Martinez Agullo¹⁶⁸, VI. Martinez Outschoorn¹⁰⁵, P. Martinez Suarez¹³,
 S. Martin-Haugh¹³⁷, G. Martinovicova¹³⁶, V.S. Martoiu^{28b}, A.C. Martyniuk⁹⁸,
 A. Marzin³⁷, D. Mascione^{79a,79b}, L. Masetti¹⁰², J. Masik¹⁰³, A.L. Maslennikov⁴⁰,
 S.L. Mason⁴³, P. Massarotti^{73a,73b}, P. Mastrandrea^{75a,75b}, A. Mastroberardino^{45b,45a},
 T. Masubuchi¹²⁷, T.T. Mathew¹²⁶, J. Matousek¹³⁶, D.M. Mattern⁵⁰, J. Maurer^{28b},
 T. Maurin⁶⁰, A.J. Maury⁶⁷, B. Maček⁹⁵, D.A. Maximov³⁹, A.E. May¹⁰³, E. Mayer⁴²,
 R. Mazini^{34g}, I. Maznas¹¹⁸, M. Mazza¹⁰⁹, S.M. Mazza¹³⁹, E. Mazzeo^{72a,72b},
 J.P. Mc Gowan¹⁷⁰, S.P. Mc Kee¹⁰⁸, C.A. Mc Lean⁶, C.C. McCracken¹⁶⁹,
 E.F. McDonald¹⁰⁷, A.E. McDougall¹¹⁷, L.F. Mcelhinney⁹³, J.A. Mcfayden¹⁵⁰,
 R.P. McGovern¹³¹, R.P. Mckenzie^{34g}, T.C. Mclachlan⁴⁹, D.J. Mclaughlin⁹⁸,
 S.J. McMahon¹³⁷, C.M. Mccpartland⁹⁴, R.A. McPherson^{170,ac}, S. Mehlhase¹¹¹,
 A. Mehta⁹⁴, D. Melini¹⁶⁸, B.R. Mellado Garcia^{34g}, A.H. Melo⁵⁶, F. Meloni⁴⁹,
 A.M. Mendes Jacques Da Costa¹⁰³, H.Y. Meng¹⁵⁹, L. Meng⁹³, S. Menke¹¹²,
 M. Mentink³⁷, E. Meoni^{45b,45a}, G. Mercado¹¹⁸, S. Merianos¹⁵⁶, C. Merlassino^{70a,70c},
 C. Meroni^{72a,72b}, J. Metcalfe⁶, A.S. Mete⁶, E. Meuser¹⁰², C. Meyer⁶⁹,
 J-P. Meyer¹³⁸, R.P. Middleton¹³⁷, L. Mijović⁵³, G. Mikenberg¹⁷⁴, M. Mikestikova¹³⁴,
 M. Mikuz⁹⁵, H. Mildner¹⁰², A. Milic³⁷, D.W. Miller⁴¹, E.H. Miller¹⁴⁷, L.S. Miller³⁵,
 A. Milov¹⁷⁴, D.A. Milstead^{48a,48b}, T. Min^{114a}, A.A. Minaenko³⁹, I.A. Minashvili^{153b},
 A.I. Mincer¹²⁰, B. Mindur^{87a}, M. Mineev⁴⁰, Y. Mino⁸⁹, L.M. Mir¹³,
 M. Miralles Lopez⁶⁰, M. Mironova^{18a}, M.C. Missio¹¹⁶, A. Mitra¹⁷², V.A. Mitsou¹⁶⁸,
 Y. Mitsumori¹¹³, O. Miu¹⁵⁹, P.S. Miyagawa⁹⁶, T. Mkrtychyan^{64a}, M. Mlinarevic⁹⁸,
 T. Mlinarevic⁹⁸, M. Mlynarikova³⁷, S. Mobius²⁰, P. Mogg¹¹¹,
 M.H. Mohamed Farook¹¹⁵, A.F. Mohammed^{14,114c}, S. Mohapatra⁴³, S. Mohiuddin¹²⁴,
 G. Mokgatitwane^{34g}, L. Moleri¹⁷⁴, B. Mondal¹⁴⁵, S. Mondal¹³⁵, K. Mönig⁴⁹,
 E. Monnier¹⁰⁴, L. Monsonis Romero¹⁶⁸, J. Montejo Berlingen¹³, A. Montella^{48a,48b},
 M. Montella¹²², F. Montereali^{78a,78b}, F. Monticelli⁹², S. Monzani^{70a,70c},
 A. Morancho Tarda⁴⁴, N. Morange⁶⁷, A.L. Moreira De Carvalho⁴⁹,

- M. Moreno Llácer¹⁶⁸, C. Moreno Martinez⁵⁷, J.M. Moreno Perez^{23b}, P. Moretini^{58b},
 S. Morgenstern³⁷, M. Morii⁶², M. Morinaga¹⁵⁷, M. Moritsu⁹⁰, F. Morodei^{76a,76b},
 P. Moschovakos³⁷, B. Moser¹²⁹, M. Mosidze^{153b}, T. Moskalets⁴⁶, P. Moskvitina¹¹⁶,
 J. Moss^{32,m}, P. Moszkowicz^{87a}, A. Moussa^{36d}, Y. Moyal¹⁷⁴, E.J.W. Moyses¹⁰⁵,
 O. Mtintsilana^{34g}, S. Muanza¹⁰⁴, J. Mueller¹³², R. Müller³⁷, G.A. Mullier¹⁶⁶,
 A.J. Mullin³³, J.J. Mullin¹³¹, A.E. Mulski⁶², D.P. Mungo¹⁵⁹, D. Munoz Perez¹⁶⁸,
 F.J. Munoz Sanchez¹⁰³, M. Murin¹⁰³, W.J. Murray^{172,137}, M. Muškinja⁹⁵,
 C. Mwewa³⁰, A.G. Myagkov^{39,a}, A.J. Myers⁸, G. Myers¹⁰⁸, M. Myska¹³⁵,
 B.P. Nachman^{18a}, K. Nagai¹²⁹, K. Nagano⁸⁵, R. Nagasaka¹⁵⁷, J.L. Nagle^{30,aj},
 E. Nagy¹⁰⁴, A.M. Nairz³⁷, Y. Nakahama⁸⁵, K. Nakamura⁸⁵, K. Nakkalil⁵,
 H. Nanjo¹²⁷, E.A. Narayanan⁴⁶, Y. Narukawa¹⁵⁷, I. Naryshkin³⁹, L. Nasella^{72a,72b},
 S. Nasri^{119b}, C. Nass²⁵, G. Navarro^{23a}, J. Navarro-Gonzalez¹⁶⁸, A. Nayaz¹⁹,
 P.Y. Nechaeva³⁹, S. Nechaeva^{24b,24a}, F. Nechansky¹³⁴, L. Nedic¹²⁹, T.J. Neep²¹,
 A. Negri^{74a,74b}, M. Negrini^{24b}, C. Nellist¹¹⁷, C. Nelson¹⁰⁶, K. Nelson¹⁰⁸,
 S. Nemecek¹³⁴, M. Nessi^{37,h}, M.S. Neubauer¹⁶⁷, F. Neuhaus¹⁰², J. Newell⁹⁴,
 P.R. Newman²¹, Y.W.Y. Ng¹⁶⁷, B. Ngair^{119a}, H.D.N. Nguyen¹¹⁰, R.B. Nickerson¹²⁹,
 R. Nicolaidou¹³⁸, J. Nielsen¹³⁹, M. Niemeyer⁵⁶, J. Niermann³⁷, N. Nikiforou³⁷,
 V. Nikolaenko^{39,a}, I. Nikolic-Audit¹³⁰, P. Nilsson³⁰, I. Ninca⁴⁹, G. Ninio¹⁵⁵,
 A. Nisati^{76a}, N. Nishu², R. Nisius¹¹², N. Nitika^{70a,70c}, J-E. Nitschke⁵¹,
 E.K. Nkadimeng^{34g}, T. Nobe¹⁵⁷, T. Nommensen¹⁵¹, M.B. Norfolk¹⁴³, B.J. Norman³⁵,
 M. Noury^{36a}, J. Novak⁹⁵, T. Novak⁹⁵, R. Novotny¹¹⁵, L. Nozka¹²⁵, K. Ntekas¹⁶³,
 N.M.J. Nunes De Moura Junior^{84b}, J. Ocariz¹³⁰, A. Ochi⁸⁶, I. Ochoa^{133a},
 S. Oerdek^{49,y}, J.T. Offermann⁴¹, A. Ogrodnik¹³⁶, A. Oh¹⁰³, C.C. Ohm¹⁴⁸,
 H. Oide⁸⁵, R. Oishi¹⁵⁷, M.L. Ojeda³⁷, Y. Okumura¹⁵⁷, L.F. Oleiro Seabra^{133a},
 I. Oleksiyuk⁵⁷, S.A. Olivares Pino^{140d}, G. Oliveira Correa¹³, D. Oliveira Damazio³⁰,
 J.L. Oliver¹⁶³, Ö.O. Öncel⁵⁵, A.P. O'Neill²⁰, A. Onofre^{133a,133e}, P.U.E. Onyisi¹¹,
 M.J. Oreglia⁴¹, D. Orestano^{78a,78b}, R.S. Orr¹⁵⁹, L.M. Osojnak¹³¹, Y. Osumi¹¹³,
 G. Otero y Garzon³¹, H. Otono⁹⁰, G.J. Ottino^{18a}, M. Ouchrif^{36d}, F. Ould-Saada¹²⁸,
 T. Ovsianikova¹⁴², M. Owen⁶⁰, R.E. Owen¹³⁷, V.E. Ozcan^{22a}, F. Ozturk⁸⁸,
 N. Ozturk⁸, S. Ozturk⁸³, H.A. Pacey¹²⁹, K. Pachal^{160a}, A. Pacheco Pages¹³,
 C. Padilla Aranda¹³, G. Padovano^{76a,76b}, S. Pagan Griso^{18a}, G. Palacino⁶⁹,
 A. Palazzo^{71a,71b}, J. Pampel²⁵, J. Pan¹⁷⁷, T. Pan^{65a}, D.K. Panchal¹¹,
 C.E. Pandini¹¹⁷, J.G. Panduro Vazquez¹³⁷, H.D. Pandya¹, H. Pang¹³⁸, P. Pani⁴⁹,
 G. Panizzo^{70a,70c}, L. Panwar¹³⁰, L. Paolozzi⁵⁷, S. Parajuli¹⁶⁷, A. Paramonov⁶,
 C. Paraskevopoulos⁵⁴, D. Paredes Hernandez^{65b}, A. Pareti^{74a,74b}, K.R. Park⁴³,
 T.H. Park¹¹², F. Parodi^{58b,58a}, J.A. Parsons⁴³, U. Parzefall⁵⁵, B. Pascual Dias⁴²,
 L. Pascual Dominguez¹⁰¹, E. Pasqualucci^{76a}, S. Passaggio^{58b}, F. Pastore⁹⁷, P. Patel⁸⁸,
 U.M. Patel⁵², J.R. Pater¹⁰³, T. Pauly³⁷, F. Pauwels¹³⁶, C.I. Pazos¹⁶², M. Pedersen¹²⁸,
 R. Pedro^{133a}, S.V. Peleganchuk³⁹, O. Penc³⁷, E.A. Pender⁵³, S. Peng¹⁵,
 G.D. Penn¹⁷⁷, K.E. Penski¹¹¹, M. Penzin³⁹, B.S. Peralva^{84d}, A.P. Pereira Peixoto¹⁴²,
 L. Pereira Sanchez¹⁴⁷, D.V. Perepelitsa^{30,aj}, G. Perera¹⁰⁵, E. Perez Codina^{160a},
 M. Perganti¹⁰, H. Pernegger³⁷, S. Perrella^{76a,76b}, O. Perrin⁴², K. Peters⁴⁹,
 R.F.Y. Peters¹⁰³, B.A. Petersen³⁷, T.C. Petersen⁴⁴, E. Petit¹⁰⁴, V. Petousis¹³⁵,
 C. Petridou^{156,e}, T. Petru¹³⁶, A. Petrukhin¹⁴⁵, M. Pettee^{18a}, A. Petukhov⁸³,
 K. Petukhova³⁷, R. Pezoa^{140f}, L. Pezzotti^{24b,24a}, G. Pezzullo¹⁷⁷, L. Pfaffenbichler³⁷,
 A.J. Pflieger³⁷, T.M. Pham¹⁷⁵, T. Pham¹⁰⁷, P.W. Phillips¹³⁷, G. Piacquadio¹⁴⁹,
 E. Pianori^{18a}, F. Piazza¹²⁶, R. Piegai³¹, D. Pietreanu^{28b}, A.D. Pilkington¹⁰³,
 M. Pinamonti^{70a,70c}, J.L. Pinfeld², B.C. Pinheiro Pereira^{133a}, J. Pinol Bel¹³,
 A.E. Pinto Pinoargote¹³⁸, L. Pintucci^{70a,70c}, K.M. Piper¹⁵⁰, A. Pirttikoski⁵⁷,
 D.A. Pizzi³⁵, L. Pizzimento^{65b}, T. Plehn^{z,an}, M.-A. Pleier³⁰, V. Pleskot¹³⁶,

E. Plotnikova⁴⁰, G. Poddar⁹⁶, R. Poettgen¹⁰⁰, L. Poggioli¹³⁰, S. Polacek¹³⁶,
G. Polesello^{74a}, A. Poley^{146,160a}, A. Polini^{24b}, C.S. Pollard¹⁷², Z.B. Pollock¹²²,
E. Pompa Pacchi¹²³, N.I. Pond⁹⁸, D. Ponomarenko⁶⁹, L. Pontecorvo³⁷, S. Popa^{28a},
G.A. Popeneciu^{28d}, A. Poreba³⁷, D.M. Portillo Quintero^{160a}, S. Pospisil¹³⁵,
M.A. Postill¹⁴³, P. Postolache^{28c}, K. Potamianos¹⁷², P.A. Potepa^{87a}, I.N. Potrap⁴⁰,
C.J. Potter³³, H. Potti¹⁵¹, J. Poveda¹⁶⁸, M.E. Pozo Astigarraga³⁷,
A. Prades Ibanez^{77a,77b}, J. Pretel¹⁷⁰, D. Price¹⁰³, M. Primavera^{71a},
L. Primomo^{70a,70c}, M.A. Principe Martin¹⁰¹, R. Privara¹²⁵, T. Procter⁶⁰,
M.L. Proffitt¹⁴², N. Proklova¹³¹, K. Prokofiev^{65c}, G. Proto¹¹², J. Proudfoot⁶,
M. Przybycien^{87a}, W.W. Przygoda^{87b}, A. Psallidas⁴⁷, J.E. Puddefoot¹⁴³, D. Pudzha⁵⁵,
D. Pyatizbyantseva¹¹⁶, J. Qian¹⁰⁸, R. Qian¹⁰⁹, D. Qichen¹⁰³, Y. Qin¹³, T. Qiu⁵³,
A. Quadt⁵⁶, M. Queitsch-Maitland¹⁰³, G. Quetant⁵⁷, R.P. Quinn¹⁶⁹,
G. Rabanal Bolanos⁶², D. Rafanoharana⁵⁵, F. Raffaelli^{77a,77b}, F. Ragusa^{72a,72b},
J.L. Rainbolt⁴¹, J.A. Raine⁵⁷, S. Rajagopalan³⁰, E. Ramakoti³⁹, L. Rambelli^{58b,58a},
I.A. Ramirez-Berend³⁵, K. Ran^{49,114c}, D.S. Rankin¹³¹, N.P. Rapheeha^{34g},
H. Rasheed^{28b}, V. Raskina¹³⁰, D.F. Rassloff^{64a}, A. Rastogi^{18a}, S. Rave¹⁰²,
S. Ravera^{58b,58a}, B. Ravina³⁷, I. Ravinovich¹⁷⁴, M. Raymond³⁷, A.L. Read¹²⁸,
N.P. Readioff¹⁴³, D.M. Rebutti^{74a,74b}, A.S. Reed¹¹², K. Reeves²⁷, J.A. Reidelsturz¹⁷⁶,
D. Reikher¹²⁶, A. Rej⁵⁰, C. Rembser³⁷, H. Ren^{63a}, M. Renda^{28b}, F. Renner⁴⁹,
A.G. Rennie¹⁶³, A.L. Rescia⁴⁹, S. Resconi^{72a}, M. Ressegotti^{58b,58a}, S. Rettie³⁷,
W.F. Rettie³⁵, J.G. Reyes Rivera¹⁰⁹, E. Reynolds^{18a}, O.L. Rezanova⁴⁰, P. Reznicek¹³⁶,
H. Riani^{36d}, N. Ribaric⁵², E. Ricci^{79a,79b}, R. Richter¹¹², S. Richter^{48a,48b},
E. Richter-Was^{87b}, M. Ridel¹³⁰, S. Ridouani^{36d}, P. Rieck¹²⁰, P. Riedler³⁷,
E.M. Riefel^{48a,48b}, J.O. Rieger¹¹⁷, M. Rijssenbeek¹⁴⁹, M. Rimoldi³⁷, L. Rinaldi^{24b,24a},
P. Rincke^{56,166}, G. Ripellino¹⁶⁶, I. Riu¹³, J.C. Rivera Vergara¹⁷⁰, F. Rizatdinova¹²⁴,
E. Rizvi⁹⁶, B.R. Roberts^{18a}, S.S. Roberts¹³⁹, D. Robinson³³, M. Robles Manzano¹⁰²,
A. Robson⁶⁰, A. Rocchi^{77a,77b}, C. Roda^{75a,75b}, S. Rodriguez Bosca³⁷,
Y. Rodriguez Garcia^{23a}, A.M. Rodríguez Vera¹¹⁸, S. Roe³⁷, J.T. Roemer³⁷, O. Røhne¹²⁸,
C.P.A. Roland¹³⁰, J. Roloff³⁰, A. Romaniouk⁸⁰, E. Romano^{74a,74b}, M. Romano^{24b},
A.C. Romero Hernandez¹⁶⁷, N. Rompotis⁹⁴, L. Roos¹³⁰, S. Rosati^{76a}, B.J. Rosser⁴¹,
E. Rossi¹²⁹, E. Rossi^{73a,73b}, L.P. Rossi⁶², L. Rossini⁵⁵, R. Rosten¹²², M. Rotaru^{28b},
B. Rottler⁵⁵, D. Rousseau⁶⁷, D. Rousso⁴⁹, S. Roy-Garand¹⁵⁹, A. Rozanov¹⁰⁴,
Z.M.A. Rozario⁶⁰, Y. Rozen¹⁵⁴, A. Rubio Jimenez¹⁶⁸, V.H. Ruelas Rivera¹⁹,
T.A. Ruggeri¹, A. Ruggiero¹²⁹, A. Ruiz-Martinez¹⁶⁸, A. Rummler³⁷, Z. Rurikova⁵⁵,
N.A. Rusakovich⁴⁰, H.L. Russell¹⁷⁰, G. Russo^{76a,76b}, J.P. Rutherford⁷,
S. Rutherford Colmenares³³, M. Rybar¹³⁶, E.B. Rye¹²⁸, A. Ryzhov⁴⁶,
J.A. Sabater Iglesias⁵⁷, H.F.W. Sadrozinski¹³⁹, F. Safai Tehrani^{76a}, S. Saha¹,
M. Sahinsoy⁸³, A. Saibel¹⁶⁸, B.T. Saifuddin¹²³, M. Saimpert¹³⁸, M. Saito¹⁵⁷,
T. Saito¹⁵⁷, A. Sala^{72a,72b}, D. Salamani³⁷, A. Salnikov¹⁴⁷, J. Salt¹⁶⁸,
A. Salvador Salas¹⁵⁵, D. Salvatore^{45b,45a}, F. Salvatore¹⁵⁰, A. Salzburger³⁷,
D. Sammel⁵⁵, E. Sampson⁹³, D. Sampsonidis^{156,e}, D. Sampsonidou¹²⁶,
J. Sánchez¹⁶⁸, V. Sanchez Sebastian¹⁶⁸, H. Sandaker¹²⁸, C.O. Sander⁴⁹,
J.A. Sandesara¹⁰⁵, M. Sandhoff¹⁷⁶, C. Sandoval^{23b}, L. Sanfilippo^{64a},
D.P.C. Sankey¹³⁷, T. Sano⁸⁹, A. Sansoni⁵⁴, L. Santi³⁷, C. Santoni⁴²,
H. Santos^{133a,133b}, A. Santra¹⁷⁴, E. Sanzani^{24b,24a}, K.A. Saoucha¹⁶⁵,
J.G. Saraiva^{133a,133d}, J. Sardain⁷, O. Sasaki⁸⁵, K. Sato¹⁶¹, C. Sauer³⁷, E. Sauvan⁴,
P. Savard^{159,ah}, R. Sawada¹⁵⁷, C. Sawyer¹³⁷, L. Sawyer⁹⁹, C. Sbarra^{24b},
A. Sbrizzi^{24b,24a}, T. Scanlon⁹⁸, J. Schaarschmidt¹⁴², U. Schäfer¹⁰²,
A.C. Schaffer^{67,46}, D. Schaile¹¹¹, R.D. Schamberger¹⁴⁹, C. Scharf¹⁹, M.M. Schefer²⁰,
V.A. Schegelsky³⁹, D. Scheirich¹³⁶, M. Schernau^{140e}, C. Scheulen⁵⁷,

- C. Schiavi^{58b,58a}, M. Schioppa^{45b,45a}, B. Schlag¹⁴⁷, S. Schlenker³⁷, J. Schmeing¹⁷⁶,
 M.A. Schmidt¹⁷⁶, K. Schmieden¹⁰², C. Schmitt¹⁰², N. Schmitt¹⁰², S. Schmitt⁴⁹,
 L. Schoeffel¹³⁸, A. Schoening^{64b}, P.G. Scholer³⁵, E. Schopf¹⁴⁵, M. Schott²⁵,
 S. Schramm⁵⁷, T. Schroer⁵⁷, H-C. Schultz-Coulon^{64a}, M. Schumacher⁵⁵,
 B.A. Schumm¹³⁹, Ph. Schune¹³⁸, H.R. Schwartz¹³⁹, A. Schwartzman¹⁴⁷,
 T.A. Schwarz¹⁰⁸, Ph. Schwemling¹³⁸, R. Schwienhorst¹⁰⁹, F.G. Sciacca²⁰,
 A. Sciandra³⁰, G. Sciolla²⁷, F. Scuri^{75a}, C.D. Sebastiani³⁷, K. Sedlaczek¹¹⁸,
 S.C. Seidel¹¹⁵, A. Seiden¹³⁹, B.D. Seidlitz⁴³, C. Seitz⁴⁹, J.M. Seixas^{84b},
 G. Sekhniaidze^{73a}, L. Selem⁶¹, N. Semprini-Cesari^{24b,24a}, A. Semushin^{178,39},
 D. Sengupta⁵⁷, V. Senthilkumar¹⁶⁸, L. Serin⁶⁷, M. Sessa^{77a,77b}, H. Severini¹²³,
 F. Sforza^{58b,58a}, A. Sfyrla⁵⁷, Q. Sha¹⁴, E. Shabalina⁵⁶, H. Shaddix¹¹⁸, A.H. Shah³³,
 R. Shaheen¹⁴⁸, J.D. Shahinian¹³¹, D. Shaked Renous¹⁷⁴, M. Shamim³⁷, L.Y. Shan¹⁴,
 M. Shapiro^{18a}, A. Sharma³⁷, A.S. Sharma¹⁶⁹, P. Sharma³⁰, P.B. Shatalov³⁹,
 K. Shaw¹⁵⁰, S.M. Shaw¹⁰³, Q. Shen^{63c}, D.J. Sheppard¹⁴⁶, P. Sherwood⁹⁸, L. Shi⁹⁸,
 X. Shi¹⁴, S. Shimizu⁸⁵, C.O. Shimmin¹⁷⁷, I.R.J. Shipsey^{129,*}, S. Shirabe⁹⁰,
 M. Shiyakova^{40,aa}, M.J. Shochet⁴¹, D.R. Shope¹²⁸, B. Shrestha¹²³, S. Shrestha^{122,al},
 I. Shreyber³⁹, M.J. Shroff¹⁷⁰, P. Sicho¹³⁴, A.M. Sickles¹⁶⁷, E. Sideras Haddad^{34g,164},
 A.C. Sidley¹¹⁷, A. Sidoti^{24b}, F. Siegert⁵¹, Dj. Sijacki¹⁶, F. Sili⁹², J.M. Silva⁵³,
 I. Silva Ferreira^{84b}, M.V. Silva Oliveira³⁰, S.B. Silverstein^{48a}, S. Simion⁶⁷,
 R. Simoniello³⁷, E.L. Simpson¹⁰³, H. Simpson¹⁵⁰, L.R. Simpson¹⁰⁸, S. Simsek⁸³,
 S. Sindhu⁵⁶, P. Sinervo¹⁵⁹, S.N. Singh²⁷, S. Singh³⁰, S. Sinha⁴⁹, S. Sinha¹⁰³,
 M. Sioli^{24b,24a}, K. Sioulas⁹, I. Siral³⁷, E. Sitnikova⁴⁹, J. Sjölin^{48a,48b}, A. Skaf⁵⁶,
 E. Skorda²¹, P. Skubic¹²³, M. Slawinska⁸⁸, I. Slazyk¹⁷, V. Smakhtin¹⁷⁴,
 B.H. Smart¹³⁷, S.Yu. Smirnov^{140b}, Y. Smirnov³⁹, L.N. Smirnova^{39,a}, O. Smirnova¹⁰⁰,
 A.C. Smith⁴³, D.R. Smith¹⁶³, E.A. Smith⁴¹, J.L. Smith¹⁰³, M.B. Smith³⁵, R. Smith¹⁴⁷,
 H. Smitmanns¹⁰², M. Smizanska⁹³, K. Smolek¹³⁵, A.A. Snesev⁴⁰, H.L. Snoek¹¹⁷,
 S. Snyder³⁰, R. Sobie^{170,ac}, A. Soffer¹⁵⁵, C.A. Solans Sanchez³⁷, E.Yu. Soldatov³⁹,
 U. Soldevila¹⁶⁸, A.A. Solodkov^{34g}, S. Solomon²⁷, A. Soloshenko⁴⁰, K. Solovieva⁵⁵,
 O.V. Solovyanov⁴², P. Sommer⁵¹, A. Sonay¹³, W.Y. Song^{160b}, A. Sopczak¹³⁵,
 A.L. Soppio⁵³, F. Sopkova^{29b}, J.D. Sorenson¹¹⁵, I.R. Sotarriva Alvarez¹⁴¹,
 V. Sothilingam^{64a}, O.J. Soto Sandoval^{140c,140b}, S. Sottocornola⁶⁹, R. Soualah¹⁶⁵,
 Z. Soumami^{36e}, D. South⁴⁹, N. Soybelman¹⁷⁴, S. Spagnolo^{71a,71b}, M. Spalla¹¹²,
 D. Sperlich⁵⁵, B. Spisso^{73a,73b}, D.P. Spiteri⁶⁰, M. Spousta¹³⁶, E.J. Staats³⁵,
 R. Stamen^{64a}, E. Stanecka⁸⁸, W. Stanek-Maslouska⁴⁹, M.V. Stange⁵¹,
 B. Stanislaus^{18a}, M.M. Stanitzki⁴⁹, B. Stapf⁴⁹, E.A. Starchenko³⁹, G.H. Stark¹³⁹,
 J. Stark⁹¹, P. Staroba¹³⁴, P. Starovoitov¹⁶⁵, R. Staszewski⁸⁸, G. Stavropoulos⁴⁷,
 A. Steffl³⁷, P. Steinberg³⁰, B. Stelzer^{146,160a}, H.J. Stelzer¹³², O. Stelzer-Chilton^{160a},
 H. Stenzel⁵⁹, T.J. Stevenson¹⁵⁰, G.A. Stewart³⁷, J.R. Stewart¹²⁴, M.C. Stockton³⁷,
 G. Stoicea^{28b}, M. Stolarski^{133a}, S. Stonjek¹¹², A. Straessner⁵¹, J. Strandberg¹⁴⁸,
 S. Strandberg^{48a,48b}, M. Stratmann¹⁷⁶, M. Strauss¹²³, T. Strebler¹⁰⁴, P. Strizenec^{29b},
 R. Ströhmer¹⁷¹, D.M. Strom¹²⁶, R. Stroynowski⁴⁶, A. Strubig^{48a,48b}, S.A. Stucci³⁰,
 B. Stugu¹⁷, J. Stupak¹²³, N.A. Styles⁴⁹, D. Su¹⁴⁷, S. Su^{63a}, W. Su^{63d}, X. Su^{63a},
 D. Suchy^{29a}, K. Sugizaki¹³¹, V.V. Sulin³⁹, M.J. Sullivan⁹⁴, D.M.S. Sultan¹²⁹,
 L. Sultanaliyeva³⁹, S. Sultansoy^{3b}, S. Sun¹⁷⁵, W. Sun¹⁴,
 O. Sunneborn Gudnadottir¹⁶⁶, N. Sur¹⁰⁴, M.R. Sutton¹⁵⁰, H. Suzuki¹⁶¹,
 M. Svatos¹³⁴, M. Swiatlowski^{160a}, T. Swirski¹⁷¹, I. Sykora^{29a}, M. Sykora¹³⁶,
 T. Sykora¹³⁶, D. Ta¹⁰², K. Tackmann^{49,y}, A. Taffard¹⁶³, R. Tafirout^{160a},
 J.S. Tafoya Vargas⁶⁷, Y. Takubo⁸⁵, M. Talby¹⁰⁴, A.A. Talyshev³⁹, K.C. Tam^{65b},
 N.M. Tamir¹⁵⁵, A. Tanaka¹⁵⁷, J. Tanaka¹⁵⁷, R. Tanaka⁶⁷, M. Tanasini¹⁴⁹, Z. Tao¹⁶⁹,
 S. Tapia Araya^{140f}, S. Tapprogge¹⁰², A. Tarek Abouelfadl Mohamed¹⁰⁹, S. Tarem¹⁵⁴,

K. Tariq¹⁴, G. Tarna^{28b}, G.F. Tartarelli^{72a}, M.J. Tartarin⁹¹, P. Tas¹³⁶,
 M. Tasevsky¹³⁴, E. Tassi^{45b,45a}, A.C. Tate¹⁶⁷, G. Tateno¹⁵⁷, Y. Tayalati^{36e,ab},
 G.N. Taylor¹⁰⁷, W. Taylor^{160b}, A.S. Tegetmeier⁹¹, P. Teixeira-Dias⁹⁷, J.J. Teoh¹⁵⁹,
 K. Terashi¹⁵⁷, J. Terron¹⁰¹, S. Terzo¹³, M. Testa⁵⁴, R.J. Teuscher^{159,ac}, A. Thaler⁸⁰,
 O. Theiner⁵⁷, T. Theveneaux-Pelzer¹⁰⁴, O. Thielmann¹⁷⁶, D.W. Thomas⁹⁷,
 J.P. Thomas²¹, E.A. Thompson^{18a}, P.D. Thompson²¹, E. Thomson¹³¹,
 R.E. Thornberry⁴⁶, C. Tian^{63a}, Y. Tian⁵⁷, V. Tikhomirov^{39,a}, Yu.A. Tikhonov³⁹,
 S. Timoshenko³⁹, D. Timoshyn¹³⁶, E.X.L. Ting¹, P. Tipton¹⁷⁷, A. Tishelman-Charny³⁰,
 S.H. Tlou^{34g}, K. Todome¹⁴¹, S. Todorova-Nova¹³⁶, S. Todt⁵¹, L. Toffolin^{70a,70c},
 M. Togawa⁸⁵, J. Tojo⁹⁰, S. Tokár^{29a}, O. Toldaiev⁶⁹, G. Tolkachev¹⁰⁴,
 M. Tomoto^{85,113}, L. Tompkins^{147,o}, E. Torrence¹²⁶, H. Torres⁹¹, E. Torró Pastor¹⁶⁸,
 M. Toscani³¹, C. Tosciri⁴¹, M. Tost¹¹, D.R. Tovey¹⁴³, T. Trefzger¹⁷¹, A. Tricoli³⁰,
 I.M. Trigger^{160a}, S. Trincaz-Duvoid¹³⁰, D.A. Trischuk²⁷, A. Tropina⁴⁰, L. Truong^{34c},
 M. Trzebinski⁸⁸, A. Trzupek⁸⁸, F. Tsai¹⁴⁹, M. Tsai¹⁰⁸, A. Tsiamis¹⁵⁶, P.V. Tsiarehka⁴⁰,
 S. Tsigaridas^{160a}, A. Tsirigotis^{156,u}, V. Tsiskaridze¹⁵⁹, E.G. Tskhadadze^{153a},
 M. Tsopoulou¹⁵⁶, Y. Tsujikawa⁸⁹, I.I. Tsukerman³⁹, V. Tsulaia^{18a}, S. Tsuno⁸⁵,
 K. Tsuru¹²¹, D. Tsybychev¹⁴⁹, Y. Tu^{65b}, A. Tudorache^{28b}, V. Tudorache^{28b},
 S. Turchikhin^{58b,58a}, I. Turk Cakir^{3a}, R. Turra^{72a}, T. Turtuvshin⁴⁰, P.M. Tuts⁴³,
 S. Tzamarias^{156,e}, E. Tzovara¹⁰², F. Ukegawa¹⁶¹, P.A. Ulloa Poblete^{140c,140b},
 E.N. Umaka³⁰, G. Unal³⁷, A. Undrus³⁰, G. Unel¹⁶³, J. Urban^{29b}, P. Urrejola^{140a},
 G. Usai⁸, R. Ushioda¹⁵⁸, M. Usman¹¹⁰, F. Ustuner⁵³, Z. Uysal⁸³, V. Vacek¹³⁵,
 B. Vachon¹⁰⁶, T. Vafeiadis³⁷, A. Vaitkus⁹⁸, C. Valderanis¹¹¹, E. Valdes Santurio^{48a,48b},
 M. Valente^{160a}, S. Valentini^{24b,24a}, A. Valero¹⁶⁸, E. Valiente Moreno¹⁶⁸,
 A. Vallier⁹¹, J.A. Valls Ferrer¹⁶⁸, D.R. Van Arneman¹¹⁷, T.R. Van Daalen¹⁴²,
 A. Van Der Graaf⁵⁰, H.Z. Van Der Schyf^{34g}, P. Van Gemmeren⁶, M. Van Rijnbach³⁷,
 S. Van Stroud⁹⁸, I. Van Vulpen¹¹⁷, P. Vana¹³⁶, M. Vanadia^{77a,77b},
 U.M. Vande Voorde¹⁴⁸, W. Vandelli³⁷, E.R. Vandewall¹²⁴, D. Vannicola¹⁵⁵,
 L. Vannoli⁵⁴, R. Vari^{76a}, E.W. Varnes⁷, C. Varni^{18b}, D. Varouchas⁶⁷, L. Varriale¹⁶⁸,
 K.E. Varvell¹⁵¹, M.E. Vasile^{28b}, L. Vaslin⁸⁵, A. Vasyukov⁴⁰, L.M. Vaughan¹²⁴,
 R. Vavricka¹³⁶, T. Vazquez Schroeder¹³, J. Veatch³², V. Vecchio¹⁰³, M.J. Veen¹⁰⁵,
 P.M. Velie^z, I. Veliscek³⁰, L.M. Veloce¹⁵⁹, F. Veloso^{133a,133c}, S. Veneziano^{76a},
 A. Ventura^{71a,71b}, S. Ventura Gonzalez¹³⁸, A. Verbytskyi¹¹², M. Verducci^{75a,75b},
 C. Vergis⁹⁶, M. Verissimo De Araujo^{84b}, W. Verkerke¹¹⁷, J.C. Vermeulen¹¹⁷,
 C. Vernieri¹⁴⁷, M. Vessella¹⁶³, M.C. Vetterli^{146,ah}, A. Vgenopoulos¹⁰²,
 N. Viaux Maira^{140f}, T. Vickey¹⁴³, O.E. Vickey Boeriu¹⁴³, G.H.A. Viehhauser¹²⁹,
 L. Vigani^{64b}, M. Vigil¹¹², M. Villa^{24b,24a}, M. Villaplana Perez¹⁶⁸, E.M. Villhauer⁵³,
 E. Vilucchi⁵⁴, M.G. Vinciter³⁵, A. Visibile¹¹⁷, C. Vittori³⁷, I. Vivarelli^{24b,24a},
 E. Voevodina¹¹², F. Vogel¹¹¹, L. Vogel^z, J.C. Voigt⁵¹, P. Vokac¹³⁵, Yu. Volkotrub^{87b},
 E. Von Toerne²⁵, B. Vormwald³⁷, K. Vorobev³⁹, M. Vos¹⁶⁸, K. Voss¹⁴⁵, M. Vozak³⁷,
 L. Vozdecky¹²³, N. Vranjes¹⁶, M. Vranjes Milosavljevic¹⁶, M. Vreeswijk¹¹⁷,
 N.K. Vu^{63d,63c}, R. Vuillermet³⁷, O. Vujanovic¹⁰², I. Vukotic⁴¹, I.K. Vyas³⁵,
 S. Wada¹⁶¹, C. Wagner¹⁴⁷, J.M. Wagner^{18a}, W. Wagner¹⁷⁶, S. Wahdan¹⁷⁶,
 H. Wahlberg⁹², C.H. Waits¹²³, J. Walder¹³⁷, R. Walker¹¹¹, W. Walkowiak¹⁴⁵,
 A. Wall¹³¹, E.J. Wallin¹⁰⁰, T. Wamorkar^{18a}, A.Z. Wang¹³⁹, C. Wang¹⁰², C. Wang¹¹,
 H. Wang^{18a}, J. Wang^{65c}, P. Wang¹⁰³, P. Wang⁹⁸, R. Wang⁶², R. Wang⁶,
 S.M. Wang¹⁵², S. Wang¹⁴, T. Wang^{63a}, W.T. Wang⁸¹, W. Wang¹⁴, X. Wang¹⁶⁷,
 X. Wang^{63c}, Y. Wang^{114a}, Y. Wang^{63a}, Z. Wang¹⁰⁸, Z. Wang^{63d,52,63c}, Z. Wang¹⁰⁸,
 C. Wanotayaroj⁸⁵, A. Warburton¹⁰⁶, R.J. Ward²¹, A.L. Warnerbring¹⁴⁵,
 N. Warrack⁶⁰, S. Waterhouse⁹⁷, A.T. Watson²¹, H. Watson⁵³, M.F. Watson²¹,
 E. Watton⁶⁰, G. Watts¹⁴², B.M. Waugh⁹⁸, J.M. Webb⁵⁵, C. Weber³⁰, H.A. Weber¹⁹,

M.S. Weber²⁰, S.M. Weber^{64a}, C. Wei^{63a}, Y. Wei⁵⁵, A.R. Weidberg¹²⁹, E.J. Weik¹²⁰,
 J. Weingarten⁵⁰, C. Weiser⁵⁵, C.J. Wells⁴⁹, T. Wenaus³⁰, B. Wendland⁵⁰,
 T. Wengler³⁷, N.S. Wenke¹¹², N. Wermes²⁵, M. Wessels^{64a}, A.M. Wharton⁹³,
 A.S. White⁶², A. White⁸, M.J. White¹, D. Whiteson¹⁶³, L. Wickremasinghe¹²⁷,
 W. Wiedenmann¹⁷⁵, M. Wielers¹³⁷, C. Wiglesworth⁴⁴, D.J. Wilbern¹²³,
 H.G. Wilkens³⁷, J.J.H. Wilkinson³³, D.M. Williams⁴³, H.H. Williams¹³¹, S. Williams³³,
 S. Willocq¹⁰⁵, B.J. Wilson¹⁰³, D.J. Wilson¹⁰³, P.J. Windischhofer⁴¹, F.I. Winkel³¹,
 F. Winklmeier¹²⁶, B.T. Winter⁵⁵, M. Wittgen¹⁴⁷, M. Wobisch⁹⁹, T. Wojtkowski⁶¹,
 Z. Wolffs¹¹⁷, J. Wollrath³⁷, M.W. Wolter⁸⁸, H. Wolters^{133a,133c}, M.C. Wong¹³⁹,
 E.L. Woodward⁴³, S.D. Worm⁴⁹, B.K. Wosiek⁸⁸, K.W. Woźniak⁸⁸, S. Wozniowski⁵⁶,
 K. Wraight⁶⁰, C. Wu²¹, M. Wu^{114b}, M. Wu¹¹⁶, S.L. Wu¹⁷⁵, X. Wu⁵⁷, X. Wu^{63a},
 Y. Wu^{63a}, Z. Wu⁴, J. Wuerzinger^{112,af}, T.R. Wyatt¹⁰³, B.M. Wynne⁵³, S. Xella⁴⁴,
 L. Xia^{114a}, M. Xia¹⁵, M. Xie^{63a}, A. Xiong¹²⁶, J. Xiong^{18a}, D. Xu¹⁴, H. Xu^{63a},
 L. Xu^{63a}, R. Xu¹³¹, T. Xu¹⁰⁸, Y. Xu¹⁴², Z. Xu⁵³, Z. Xu^{114a}, B. Yabsley¹⁵¹,
 S. Yacoob^{34a}, Y. Yamaguchi⁸⁵, E. Yamashita¹⁵⁷, H. Yamauchi¹⁶¹, T. Yamazaki^{18a},
 Y. Yamazaki⁸⁶, S. Yan⁶⁰, Z. Yan¹⁰⁵, H.J. Yang^{63c,63d}, H.T. Yang^{63a}, S. Yang^{63a},
 T. Yang^{65c}, X. Yang³⁷, X. Yang¹⁴, Y. Yang⁴⁶, Y. Yang^{63a}, W-M. Yao^{18a}, H. Ye⁵⁶,
 J. Ye¹⁴, S. Ye³⁰, X. Ye^{63a}, Y. Yeh⁹⁸, I. Yeletsikh⁴⁰, B. Yeo^{18b}, M.R. Yexley⁹⁸,
 T.P. Yildirim¹²⁹, P. Yin⁴³, K. Yorita¹⁷³, S. Younas^{28b}, C.J.S. Young³⁷, C. Young¹⁴⁷,
 N.D. Young¹²⁶, Y. Yu^{63a}, J. Yuan^{14,114c}, M. Yuan¹⁰⁸, R. Yuan^{63d,63c}, L. Yue⁹⁸,
 M. Zaazoua^{63a}, B. Zabinski⁸⁸, I. Zahir^{36a}, Z.K. Zak⁸⁸, T. Zakareishvili¹⁶⁸,
 S. Zambito⁵⁷, J.A. Zamora Saa^{140d,140b}, J. Zang¹⁵⁷, D. Zanzi⁵⁵, R. Zanzottera^{72a,72b},
 O. Zaplatilek¹³⁵, C. Zeitnitz¹⁷⁶, H. Zeng¹⁴, J.C. Zeng¹⁶⁷, D.T. Zenger Jr²⁷,
 O. Zenin³⁹, T. Ženiš^{29a}, S. Zenz⁹⁶, S. Zerradi^{36a}, D. Zerwas⁶⁷, M. Zhai^{14,114c},
 D.F. Zhang¹⁴³, J. Zhang^{63b}, J. Zhang⁶, K. Zhang^{14,114c}, L. Zhang^{63a}, L. Zhang^{114a},
 P. Zhang^{14,114c}, R. Zhang¹⁷⁵, S. Zhang⁹¹, T. Zhang¹⁵⁷, X. Zhang^{63c}, Y. Zhang¹⁴²,
 Y. Zhang⁹⁸, Y. Zhang^{63a}, Y. Zhang^{114a}, Z. Zhang^{18a}, Z. Zhang^{63b}, Z. Zhang⁶⁷,
 H. Zhao¹⁴², T. Zhao^{63b}, Y. Zhao³⁵, Z. Zhao^{63a}, Z. Zhao^{63a}, A. Zhemchugov⁴⁰,
 J. Zheng^{114a}, K. Zheng¹⁶⁷, X. Zheng^{63a}, Z. Zheng¹⁴⁷, D. Zhong¹⁶⁷, B. Zhou¹⁰⁸,
 H. Zhou⁷, N. Zhou^{63c}, Y. Zhou¹⁵, Y. Zhou^{114a}, Y. Zhou⁷, C.G. Zhu^{63b}, J. Zhu¹⁰⁸,
 X. Zhu^{63d}, Y. Zhu^{63c}, Y. Zhu^{63a}, X. Zhuang¹⁴, K. Zhukov⁶⁹, N.I. Zimine⁴⁰,
 J. Zinsser^{64b}, M. Ziolkowski¹⁴⁵, L. Živković¹⁶, A. Zoccoli^{24b,24a}, K. Zoch⁶²,
 T.G. Zorbas¹⁴³, O. Zormpa⁴⁷, W. Zou⁴³, L. Zwalinski³⁷

¹Department of Physics, University of Adelaide, Adelaide; Australia

²Department of Physics, University of Alberta, Edmonton AB; Canada

^{3(a)}Department of Physics, Ankara University, Ankara; ^(b)Division of Physics,
TOBB University of Economics and Technology, Ankara; Türkiye

⁴LAPP, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy; France

⁵APC, Université Paris Cité, CNRS/IN2P3, Paris; France

⁶High Energy Physics Division, Argonne National Laboratory, Argonne IL;
United States of America

⁷Department of Physics, University of Arizona, Tucson AZ; United States of America

⁸Department of Physics, University of Texas at Arlington, Arlington TX;
United States of America

⁹Physics Department, National and Kapodistrian University of Athens, Athens; Greece

¹⁰Physics Department, National Technical University of Athens, Zografou; Greece

¹¹Department of Physics, University of Texas at Austin, Austin TX;
United States of America

¹²Institute of Physics, Azerbaijan Academy of Sciences, Baku; Azerbaijan

- ¹³Institut de Física d'Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona; Spain
- ¹⁴Institute of High Energy Physics, Chinese Academy of Sciences, Beijing; China
- ¹⁵Physics Department, Tsinghua University, Beijing; China
- ¹⁶Institute of Physics, University of Belgrade, Belgrade; Serbia
- ¹⁷Department for Physics and Technology, University of Bergen, Bergen; Norway
- ^{18(a)}Physics Division, Lawrence Berkeley National Laboratory, Berkeley CA; ^(b)University of California, Berkeley CA; United States of America
- ¹⁹Institut für Physik, Humboldt Universität zu Berlin, Berlin; Germany
- ²⁰Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern; Switzerland
- ²¹School of Physics and Astronomy, University of Birmingham, Birmingham; United Kingdom
- ^{22(a)}Department of Physics, Bogazici University, Istanbul; ^(b)Department of Physics Engineering, Gaziantep University, Gaziantep; ^(c)Department of Physics, Istanbul University, Istanbul; Türkiye
- ^{23(a)}Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá; ^(b)Departamento de Física, Universidad Nacional de Colombia, Bogotá; Colombia
- ^{24(a)}Dipartimento di Fisica e Astronomia A. Righi, Università di Bologna, Bologna; ^(b)INFN Sezione di Bologna; Italy
- ²⁵Physikalisches Institut, Universität Bonn, Bonn; Germany
- ²⁶Department of Physics, Boston University, Boston MA; United States of America
- ²⁷Department of Physics, Brandeis University, Waltham MA; United States of America
- ^{28(a)}Transilvania University of Brasov, Brasov; ^(b)Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest; ^(c)Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi; ^(d)National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj-Napoca; ^(e)National University of Science and Technology Politehnica, Bucharest; ^(f)West University in Timisoara, Timisoara; ^(g)Faculty of Physics, University of Bucharest, Bucharest; Romania
- ^{29(a)}Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava; ^(b)Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice; Slovak Republic
- ³⁰Physics Department, Brookhaven National Laboratory, Upton NY; United States of America
- ³¹Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física, y CONICET, Instituto de Física de Buenos Aires (IFIBA), Buenos Aires; Argentina
- ³²California State University, CA; United States of America
- ³³Cavendish Laboratory, University of Cambridge, Cambridge; United Kingdom
- ^{34(a)}Department of Physics, University of Cape Town, Cape Town; ^(b)iThemba Labs, Western Cape; ^(c)Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg; ^(d)National Institute of Physics, University of the Philippines Diliman (Philippines); ^(e)University of South Africa, Department of Physics, Pretoria; ^(f)University of Zululand, KwaDlangezwa; ^(g)School of Physics, University of the Witwatersrand, Johannesburg; South Africa
- ³⁵Department of Physics, Carleton University, Ottawa ON; Canada
- ^{36(a)}Faculté des Sciences Ain Chock, Université Hassan II de Casablanca; ^(b)Faculté des Sciences, Université Ibn-Tofail, Kénitra; ^(c)Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech; ^(d)LPMR, Faculté des Sciences, Université Mohamed Premier, Oujda; ^(e)Faculté des sciences, Université Mohammed V, Rabat; ^(f)Institute of Applied Physics, Mohammed VI Polytechnic University, Ben Guerir; Morocco
- ³⁷CERN, Geneva; Switzerland
- ³⁸Affiliated with an institute formerly covered by a cooperation agreement with CERN

- ³⁹Affiliated with an institute covered by a cooperation agreement with CERN
- ⁴⁰Affiliated with an international laboratory covered by a cooperation agreement with CERN
- ⁴¹Enrico Fermi Institute, University of Chicago, Chicago IL; United States of America
- ⁴²LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand; France
- ⁴³Nevis Laboratory, Columbia University, Irvington NY; United States of America
- ⁴⁴Niels Bohr Institute, University of Copenhagen, Copenhagen; Denmark
- ^{45(a)}Dipartimento di Fisica, Università della Calabria, Rende;
- ^(b)INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati; Italy
- ⁴⁶Physics Department, Southern Methodist University, Dallas TX; United States of America
- ⁴⁷National Centre for Scientific Research "Demokritos", Agia Paraskevi; Greece
- ^{48(a)}Department of Physics, Stockholm University; ^(b)Oskar Klein Centre, Stockholm; Sweden
- ⁴⁹Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen; Germany
- ⁵⁰Fakultät Physik, Technische Universität Dortmund, Dortmund; Germany
- ⁵¹Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden; Germany
- ⁵²Department of Physics, Duke University, Durham NC; United States of America
- ⁵³SUPA - School of Physics and Astronomy, University of Edinburgh, Edinburgh; United Kingdom
- ⁵⁴INFN e Laboratori Nazionali di Frascati, Frascati; Italy
- ⁵⁵Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg; Germany
- ⁵⁶II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen; Germany
- ⁵⁷Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
- ^{58(a)}Dipartimento di Fisica, Università di Genova, Genova; ^(b)INFN Sezione di Genova; Italy
- ⁵⁹II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen; Germany
- ⁶⁰SUPA - School of Physics and Astronomy, University of Glasgow, Glasgow; United Kingdom
- ⁶¹LPSC, Université Grenoble Alpes, CNRS/IN2P3, Grenoble INP, Grenoble; France
- ⁶²Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge MA; United States of America
- ^{63(a)}Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei; ^(b)Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao; ^(c)School of Physics and Astronomy, Shanghai Jiao Tong University, Key Laboratory for Particle Astrophysics and Cosmology (MOE), SKLPPC, Shanghai; ^(d)Tsung-Dao Lee Institute, Shanghai;
- ^(e)School of Physics, Zhengzhou University; China
- ^{64(a)}Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg; ^(b)Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg; Germany
- ^{65(a)}Department of Physics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong; ^(b)Department of Physics, University of Hong Kong, Hong Kong; ^(c)Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; China
- ⁶⁶Department of Physics, National Tsing Hua University, Hsinchu; Taiwan
- ⁶⁷IJCLab, Université Paris-Saclay, CNRS/IN2P3, 91405, Orsay; France
- ⁶⁸Centro Nacional de Microelectrónica (IMB-CNM-CSIC), Barcelona; Spain
- ⁶⁹Department of Physics, Indiana University, Bloomington IN; United States of America

- ^{70(a)}INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine;^(b)ICTP, Trieste;^(c)Dipartimento Politecnico di Ingegneria e Architettura, Università di Udine, Udine; Italy
- ^{71(a)}INFN Sezione di Lecce;^(b)Dipartimento di Matematica e Fisica, Università del Salento, Lecce; Italy
- ^{72(a)}INFN Sezione di Milano;^(b)Dipartimento di Fisica, Università di Milano, Milano; Italy
- ^{73(a)}INFN Sezione di Napoli;^(b)Dipartimento di Fisica, Università di Napoli, Napoli; Italy
- ^{74(a)}INFN Sezione di Pavia;^(b)Dipartimento di Fisica, Università di Pavia, Pavia; Italy
- ^{75(a)}INFN Sezione di Pisa;^(b)Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa; Italy
- ^{76(a)}INFN Sezione di Roma;^(b)Dipartimento di Fisica, Sapienza Università di Roma, Roma; Italy
- ^{77(a)}INFN Sezione di Roma Tor Vergata;^(b)Dipartimento di Fisica, Università di Roma Tor Vergata, Roma; Italy
- ^{78(a)}INFN Sezione di Roma Tre;^(b)Dipartimento di Matematica e Fisica, Università Roma Tre, Roma; Italy
- ^{79(a)}INFN-TIFPA;^(b)Università degli Studi di Trento, Trento; Italy
- ⁸⁰Universität Innsbruck, Department of Astro and Particle Physics, Innsbruck; Austria
- ⁸¹University of Iowa, Iowa City IA; United States of America
- ⁸²Department of Physics and Astronomy, Iowa State University, Ames IA; United States of America
- ⁸³Istinye University, Sariyer, Istanbul; Türkiye
- ^{84(a)}Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora;^(b)Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro;^(c)Instituto de Física, Universidade de São Paulo, São Paulo;^(d)Rio de Janeiro State University, Rio de Janeiro;^(e)Federal University of Bahia, Bahia; Brazil
- ⁸⁵KEK, High Energy Accelerator Research Organization, Tsukuba; Japan
- ⁸⁶Graduate School of Science, Kobe University, Kobe; Japan
- ^{87(a)}AGH University of Krakow, Faculty of Physics and Applied Computer Science, Krakow;^(b)Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow; Poland
- ⁸⁸Institute of Nuclear Physics Polish Academy of Sciences, Krakow; Poland
- ⁸⁹Faculty of Science, Kyoto University, Kyoto; Japan
- ⁹⁰Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka; Japan
- ⁹¹L2IT, Université de Toulouse, CNRS/IN2P3, UPS, Toulouse; France
- ⁹²Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata; Argentina
- ⁹³Physics Department, Lancaster University, Lancaster; United Kingdom
- ⁹⁴Oliver Lodge Laboratory, University of Liverpool, Liverpool; United Kingdom
- ⁹⁵Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana; Slovenia
- ⁹⁶School of Physics and Astronomy, Queen Mary University of London, London; United Kingdom
- ⁹⁷Department of Physics, Royal Holloway University of London, Egham; United Kingdom
- ⁹⁸Department of Physics and Astronomy, University College London, London; United Kingdom
- ⁹⁹Louisiana Tech University, Ruston LA; United States of America
- ¹⁰⁰Fysiska institutionen, Lunds universitet, Lund; Sweden
- ¹⁰¹Departamento de Física Teórica C-15 and CIAFF, Universidad Autónoma de Madrid, Madrid; Spain
- ¹⁰²Institut für Physik, Universität Mainz, Mainz; Germany

- ¹⁰³School of Physics and Astronomy, University of Manchester, Manchester; United Kingdom
- ¹⁰⁴CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille; France
- ¹⁰⁵Department of Physics, University of Massachusetts, Amherst MA; United States of America
- ¹⁰⁶Department of Physics, McGill University, Montreal QC; Canada
- ¹⁰⁷School of Physics, University of Melbourne, Victoria; Australia
- ¹⁰⁸Department of Physics, University of Michigan, Ann Arbor MI; United States of America
- ¹⁰⁹Department of Physics and Astronomy, Michigan State University, East Lansing MI; United States of America
- ¹¹⁰Group of Particle Physics, University of Montreal, Montreal QC; Canada
- ¹¹¹Fakultät für Physik, Ludwig-Maximilians-Universität München, München; Germany
- ¹¹²Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München; Germany
- ¹¹³Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya; Japan
- ^{114(a)}Department of Physics, Nanjing University, Nanjing; ^(b)School of Science, Shenzhen Campus of Sun Yat-sen University; ^(c)University of Chinese Academy of Science (UCAS), Beijing; China
- ¹¹⁵Department of Physics and Astronomy, University of New Mexico, Albuquerque NM; United States of America
- ¹¹⁶Institute for Mathematics, Astrophysics and Particle Physics, Radboud University/Nikhef, Nijmegen; Netherlands
- ¹¹⁷Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam; Netherlands
- ¹¹⁸Department of Physics, Northern Illinois University, DeKalb IL; United States of America
- ^{119(a)}New York University Abu Dhabi, Abu Dhabi; ^(b)United Arab Emirates University, Al Ain; United Arab Emirates
- ¹²⁰Department of Physics, New York University, New York NY; United States of America
- ¹²¹Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo; Japan
- ¹²²Ohio State University, Columbus OH; United States of America
- ¹²³Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman OK; United States of America
- ¹²⁴Department of Physics, Oklahoma State University, Stillwater OK; United States of America
- ¹²⁵Palacký University, Joint Laboratory of Optics, Olomouc; Czech Republic
- ¹²⁶Institute for Fundamental Science, University of Oregon, Eugene, OR; United States of America
- ¹²⁷Graduate School of Science, Osaka University, Osaka; Japan
- ¹²⁸Department of Physics, University of Oslo, Oslo; Norway
- ¹²⁹Department of Physics, Oxford University, Oxford; United Kingdom
- ¹³⁰LPNHE, Sorbonne Université, Université Paris Cité, CNRS/IN2P3, Paris; France
- ¹³¹Department of Physics, University of Pennsylvania, Philadelphia PA; United States of America
- ¹³²Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh PA; United States of America
- ^{133(a)}Laboratório de Instrumentação e Física Experimental de Partículas - LIP, Lisboa; ^(b)Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Lisboa; ^(c)Departamento de Física, Universidade de Coimbra, Coimbra; ^(d)Centro de Física Nuclear da Universidade de Lisboa, Lisboa; ^(e)Departamento de Física, Universidade do Minho, Braga; ^(f)Departamento de Física Teórica y del Cosmos, Universidad de Granada, Granada (Spain); ^(g)Departamento de Física,

- Instituto Superior Técnico, Universidade de Lisboa, Lisboa; Portugal
- ¹³⁴Institute of Physics of the Czech Academy of Sciences, Prague; Czech Republic
- ¹³⁵Czech Technical University in Prague, Prague; Czech Republic
- ¹³⁶Charles University, Faculty of Mathematics and Physics, Prague; Czech Republic
- ¹³⁷Particle Physics Department, Rutherford Appleton Laboratory, Didcot; United Kingdom
- ¹³⁸IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette; France
- ¹³⁹Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz CA; United States of America
- ^{140(a)}Departamento de Física, Pontificia Universidad Católica de Chile, Santiago;^(b)Millennium Institute for Subatomic physics at high energy frontier (SAPHIR), Santiago;^(c)Instituto de Investigación Multidisciplinario en Ciencia y Tecnología, y Departamento de Física, Universidad de La Serena;^(d)Universidad Andres Bello, Department of Physics, Santiago;^(e)Instituto de Alta Investigación, Universidad de Tarapacá, Arica;^(f)Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso; Chile
- ¹⁴¹Department of Physics, Institute of Science, Tokyo; Japan
- ¹⁴²Department of Physics, University of Washington, Seattle WA; United States of America
- ¹⁴³Department of Physics and Astronomy, University of Sheffield, Sheffield; United Kingdom
- ¹⁴⁴Department of Physics, Shinshu University, Nagano; Japan
- ¹⁴⁵Department Physik, Universität Siegen, Siegen; Germany
- ¹⁴⁶Department of Physics, Simon Fraser University, Burnaby BC; Canada
- ¹⁴⁷SLAC National Accelerator Laboratory, Stanford CA; United States of America
- ¹⁴⁸Department of Physics, Royal Institute of Technology, Stockholm; Sweden
- ¹⁴⁹Departments of Physics and Astronomy, Stony Brook University, Stony Brook NY; United States of America
- ¹⁵⁰Department of Physics and Astronomy, University of Sussex, Brighton; United Kingdom
- ¹⁵¹School of Physics, University of Sydney, Sydney; Australia
- ¹⁵²Institute of Physics, Academia Sinica, Taipei; Taiwan
- ^{153(a)}E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi;^(b)High Energy Physics Institute, Tbilisi State University, Tbilisi;^(c)University of Georgia, Tbilisi; Georgia
- ¹⁵⁴Department of Physics, Technion, Israel Institute of Technology, Haifa; Israel
- ¹⁵⁵Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv; Israel
- ¹⁵⁶Department of Physics, Aristotle University of Thessaloniki, Thessaloniki; Greece
- ¹⁵⁷International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo; Japan
- ¹⁵⁸Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo; Japan
- ¹⁵⁹Department of Physics, University of Toronto, Toronto ON; Canada
- ^{160(a)}TRIUMF, Vancouver BC;^(b)Department of Physics and Astronomy, York University, Toronto ON; Canada
- ¹⁶¹Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba; Japan
- ¹⁶²Department of Physics and Astronomy, Tufts University, Medford MA; United States of America
- ¹⁶³Department of Physics and Astronomy, University of California Irvine, Irvine CA; United States of America
- ¹⁶⁴University of West Attica, Athens; Greece
- ¹⁶⁵University of Sharjah, Sharjah; United Arab Emirates
- ¹⁶⁶Department of Physics and Astronomy, University of Uppsala, Uppsala; Sweden
- ¹⁶⁷Department of Physics, University of Illinois, Urbana IL; United States of America

- ¹⁶⁸Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia - CSIC, Valencia; Spain
- ¹⁶⁹Department of Physics, University of British Columbia, Vancouver BC; Canada
- ¹⁷⁰Department of Physics and Astronomy, University of Victoria, Victoria BC; Canada
- ¹⁷¹Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg; Germany
- ¹⁷²Department of Physics, University of Warwick, Coventry; United Kingdom
- ¹⁷³Waseda University, Tokyo; Japan
- ¹⁷⁴Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot; Israel
- ¹⁷⁵Department of Physics, University of Wisconsin, Madison WI; United States of America
- ¹⁷⁶Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal; Germany
- ¹⁷⁷Department of Physics, Yale University, New Haven CT; United States of America
- ¹⁷⁸Yerevan Physics Institute, Yerevan; Armenia
- ^aAlso Affiliated with an institute covered by a cooperation agreement with CERN
- ^bAlso at An-Najah National University, Nablus; Palestine
- ^cAlso at Borough of Manhattan Community College, City University of New York, New York NY; United States of America
- ^dAlso at Center for High Energy Physics, Peking University; China
- ^eAlso at Center for Interdisciplinary Research and Innovation (CIRI-AUTH), Thessaloniki; Greece
- ^fAlso at CERN, Geneva; Switzerland
- ^gAlso at CMD-AC UNEC Research Center, Azerbaijan State University of Economics (UNEC); Azerbaijan
- ^hAlso at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
- ⁱAlso at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona; Spain
- ^jAlso at Department of Financial and Management Engineering, University of the Aegean, Chios; Greece
- ^kAlso at Department of Mathematical Sciences, University of South Africa, Johannesburg; South Africa
- ^lAlso at Department of Physics, Bolu Abant İzzet Baysal University, Bolu; Türkiye
- ^mAlso at Department of Physics, California State University, Sacramento; United States of America
- ⁿAlso at Department of Physics, King's College London, London; United Kingdom
- ^oAlso at Department of Physics, Stanford University, Stanford CA; United States of America
- ^pAlso at Department of Physics, Stellenbosch University; South Africa
- ^qAlso at Department of Physics, University of Fribourg, Fribourg; Switzerland
- ^rAlso at Department of Physics, University of Thessaly; Greece
- ^sAlso at Department of Physics, Westmont College, Santa Barbara; United States of America
- ^tAlso at Faculty of Physics, Sofia University, 'St. Kliment Ohridski', Sofia; Bulgaria
- ^uAlso at Hellenic Open University, Patras; Greece
- ^vAlso at Henan University; China
- ^wAlso at Imam Mohammad Ibn Saud Islamic University; Saudi Arabia
- ^xAlso at Institut Catalana de Recerca i Estudis Avançats, ICREA, Barcelona; Spain
- ^yAlso at Institut für Experimentalphysik, Universität Hamburg, Hamburg; Germany
- ^zAssociated at Institut für Theoretische Physik, Universität Heidelberg, Heidelberg; Germany

^{aa}Also at Institute for Nuclear Research and Nuclear Energy (INRNE)
of the Bulgarian Academy of Sciences, Sofia; Bulgaria

^{ab}Also at Institute of Applied Physics, Mohammed VI Polytechnic University,
Ben Guerir; Morocco

^{ac}Also at Institute of Particle Physics (IPP); Canada

^{ad}Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku; Azerbaijan

^{ae}Also at National Institute of Physics, University of the Philippines Diliman (Philippines);
Philippines

^{af}Also at Technical University of Munich, Munich; Germany

^{ag}Also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing; China

^{ah}Also at TRIUMF, Vancouver BC; Canada

^{ai}Also at Università di Napoli Parthenope, Napoli; Italy

^{aj}Also at University of Colorado Boulder, Department of Physics,
Colorado; United States of America

^{ak}Also at University of the Western Cape; South Africa

^{al}Also at Washington College, Chestertown, MD; United States of America

^{am}Also at Yeditepe University, Physics Department, Istanbul; Türkiye

^{an}Also at Institute for Scientific Computing (IWR), University of Heidelberg,
Heidelberg; Germany

* Deceased