

Machine Learning Classification of Sphalerons and Black Holes at the LHC

Aurora Singstad Grefsrud^(a), Trygve Buanes^(a), Fotis Koutroulis^(c),
Anna Lipniacka^(b), Rafał Maselek^{(c),(e)}, Andreas Papaefstathiou^(d), Kazuki Sakurai^(c),
Therese B. Sjursen^(a) and Igor Slazyk^(a)

^(a)*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western
Norway University of Applied Sciences,
Postbox 7030, 5020 Bergen, Norway*

^(b)*Department of Physics and Technology, University of Bergen,
Postboks 7803, N-5020 Bergen, Norway*

^(c)*Institute of Theoretical Physics, Faculty of Physics,
University of Warsaw, ul. Pasteura 5, PL-02-093 Warsaw, Poland*

^(d)*Department of Physics, Kennesaw State University,
830 Polytechnic Lane, Marietta, GA 30060, USA*

^(e)*Laboratoire de Physique Subatomique et de Cosmologie (LPSC), Université Grenoble-Alpes,
CNRS/IN2P3, 53 Avenue des Martyrs, F-38026 Grenoble, France*

Abstract

In models with large extra dimensions, “miniature” black holes (BHs) might be produced in high-energy proton-proton collisions at the Large Hadron Collider (LHC). In the semi-classical regime, those BHs thermally decay, giving rise to large-multiplicity final states with jets and leptons. On the other hand, similar final states are also expected in the production of electroweak sphaleron/instanton-induced processes. We investigate whether one can discriminate these scenarios when BH or sphaleron-like events are observed in the LHC using Machine Learning (ML) methods. Classification among several BH scenarios with different numbers of extra dimensions and the minimal BH masses is also examined. In this study we consider three ML models: XGBoost algorithms with (1) high- and (2) low-level inputs, and (3) a Residual Convolutional Neural Network. In the latter case, the low-level detector information is converted into an input format of three-layer binned event images, where the value of each bin corresponds to the energy deposited in various detector subsystems. We demonstrate that only a few detected events are sufficient to effectively discriminate between the sphaleron and BH processes. Separation among BH scenarios with different minimal BH masses is also possible with a reasonable number of events, that can be collected in the LHC Run-2, -3 and the high-luminosity LHC (HL-LHC). We find, however, that a large number of events is needed to discriminate between BH hypotheses with the same minimal BH mass, but different numbers of extra dimensions.

1 Introduction

The Standard Model (SM) of particle physics has been extremely successful in describing particle interactions below the TeV scale. On the other hand, the SM has several theoretical issues, such as the hierarchy problem, as well as phenomenological problems, for instance, the inability to account for the existence of dark matter, and the observed asymmetry between matter and antimatter in the Universe. These problems suggest that a more fundamental theory underlying the SM must exist, and may emerge at the energy scale around, or higher, than the TeV scale.

An attractive solution to the hierarchy problem is the scenario with *Large Extra Dimensions* (LEDs) [1–3]. This model postulates that all matter and gauge fields, except for gravity, are restricted to live on a $(3 + 1)$ -dimensional hypersurface, called a 3-brane, which is embedded in the higher dimensional spacetime. Assuming all extra dimensions, orthogonal to the 3-brane, are compactified, the traditional Planck scale $M_P \sim 10^{18}$ GeV is understood as an effective scale, derived from the fundamental higher-dimensional Planck scale M_* with the relation:

$$M_*^{2+n} \sim \frac{M_P^2}{R^n}, \quad (1.1)$$

where we have assumed the existence of n extra dimensions, with the common size R . If the size of extra dimensions is much larger than the traditional Planck length $R \gg 1/M_P$, the fundamental Planck scale M_* can be much smaller than the usual Planck scale M_P . In particular, the hierarchy problem is solved if M_* is around the TeV scale [1–3].

This scenario has a striking implication for new physics searches at high-energy colliders, such as the Large Hadron Collider (LHC) at CERN. For the fundamental Planck scale around or smaller than the TeV scale, $M_* \lesssim \mathcal{O}(1)$ TeV, high-energy colliders can offer trans-Planckian particle collisions with $\sqrt{s} > M_*$. In this situation, the gravitational interaction becomes very strong and the colliding particles may collapse into a black hole (BH) [4–11]. Such a “miniature” BH, once formed, will quickly “evaporate” through Hawking radiation [12]. The spectrum of emitted particles can be understood as the thermal radiation characterized by the Hawking temperature T_H of the BH at a given stage of the evaporation [13, 14]. Those outgoing particles, however, must travel through a strong gravitational potential created by the BH, and the spectrum is distorted from the blackbody profile for an observer located infinitely far from the BH [15, 16]. The deviation from the blackbody spectrum is described by the so-called “greybody factor”, which carries the information of extra dimensions [14].

The last observation poses an interesting question to collider physics: “can we discriminate one extra dimension scenario from another with the BH signature at a collider, if observed?”, or more concretely, for example, “can we identify the number of extra dimensions by analysing BH events?”. Since the greybody factor depends on the parameters of extra dimensions in a subtle way, answering these questions is non-trivial.

There even exists a similar collider signature that has a completely different origin, due to electroweak (EW) sphaleron/instanton-induced processes [17–20]. These are non-perturbative processes within the SM, which go over (or penetrate via quantum tunneling) a potential barrier separating two distinctive EW vacua: the Higgs and EW gauge field configurations in these vacua are characterised by different values of the topological winding number, N_{CS} (the Chern-Simons number). The EW sphaleron plays a crucial role in many proposed scenarios of baryogenesis (see, e.g., Ref. [21] for a review). The theoretical prediction for the production cross-section of these non-perturbative processes at high-energy colliders suffers from large uncertainties, and whether such processes are observable in the foreseeable future is still under debate [22–39]. Although the cross-section is largely unknown, the final state of the process is anticipated to possess certain

characteristic features. Firstly, the minimum height of the potential barrier, E_{sph} , is known to good accuracy, $E_{\text{sph}} \simeq 9.1$ TeV [17, 18]. Therefore, one expects that those processes can occur only at very high energies, at least with $\sqrt{\hat{s}} \gtrsim E_{\text{sph}}$. Secondly, since the anomaly connects the change of N_{CS} , and the change of the fermion number that couples to the $SU(2)_L$ gauge bosons, *all* left-handed fermions, $(3 \text{ quarks} + 1 \text{ lepton}) \times (3 \text{ generations})$, of the SM must be involved in the interaction [40, 41]. Therefore, ‘consuming’ two light-flavor quarks in the initial state, the final state must contain at least seven anti-quarks and three anti-leptons (plus some EW bosons). Such a high-energy and high-multiplicity final state resembles the BH events mentioned. It is therefore a non-trivial task to discriminate the EW sphaleron/instanton-induced processes from the miniature BH signature in the LED scenario.

A traditional way of tackling the above questions is to analyse various collider variables built out of reconstructed objects, such as jets, isolated leptons, and the missing transverse energy, E_T^{miss} . In fact, as we will show later, the jet multiplicity and the charge asymmetry of muons turn out to be powerful discriminators between the BH and the EW sphaleron/instanton signatures. On the other hand, this approach has several drawbacks. In the aforementioned signal events (BHs and EW sphalerons/instantons), the final-state multiplicity is very large and jets and leptons often overlap with each other. In such a busy environment, the multiplicity of reconstructed objects is not a robust quantity, in the sense that it is sensitive to the details of the chosen jet definition and isolation criteria. Moreover, muons are often rejected by the isolation condition. As a result, the majority of signal events contain no reconstructed muons. Lastly, the final-state multiplicity and the muon charge asymmetry are not useful discriminators to identify the number of extra dimensions with the BH signature.

The identification of the number of extra dimensions, and the discrimination between the BH and EW sphaleron/instanton events, is essentially a classification problem, which machine learning (ML) methods have been proven to be exceptionally good at. The application of ML to collider physics has also been an active research field in the last decade. Recent development includes application of deep learning, i.e. neural networks (NNs) to, for example, triggering, background estimation, jet tagging, and event classification (see, e.g., Ref. [42] for a comprehensive list of references). More recently, deep learning methods with low-level data of the whole event (e.g. the energy deposits of the entire calorimeter) have been investigated [43–45]. These studies represented entire collider events as images, and processed them using Convolutional Neural Networks (CNNs) [46]. They have demonstrated that such methods can exceed the sensitivity of standard approaches with high-level inputs, such as the four-momenta of reconstructed objects. Classifying events using low-level inputs has an obvious advantage in our problem, since the object reconstruction process is entirely omitted, and all detector information is kept and used in the analysis. It should be noted, however, that this approach has certain limitations in high pile-up conditions, when the detector is swamped by energy deposits that do not belong to the main hard interaction of interest. Reconstructions of jets pointing to the main vertex of interest, and rejecting the deposits not belonging to them, typically serves as pile-up mitigation. In what follows, we assume that pile-up mitigation can be performed, and we do not consider energy deposits coming from pile-up interactions. Furthermore, we assume that the SM backgrounds can be suppressed with the CMS-inspired selections [47], which are applied to our data.

In this paper, we study the discrimination of EW sphaleron and five different BH scenarios, using three different ML methods. The aim is to see if we can separate these scenarios with a reasonable number of events that can be collected in the on-going and future LHC runs. Three ML methods are examined. The first two are based on the state-of-the-art XGBoost library, which constructs a decision tree model. The low-level tracking and calorimeter data are used as inputs in the first method, whereas the second method employs tabular reconstructed object data. The third

Hypothesis	n	M_{\min}	$E_{\text{thr}}^{\text{sph}}$	ϵ_{sig}
SPH_9	—	—	9 TeV	0.18
BH_n4_M8	4	8 TeV	—	0.17
BH_n2_M10	2	10 TeV	—	0.55
BH_n4_M10	4	10 TeV	—	0.49
BH_n6_M10	6	10 TeV	—	0.43
BH_n4_M12	4	12 TeV	—	0.65

Table 1: List of the six model hypotheses used in this study. The first row is the sphaleron scenario with the 9 TeV threshold energy, $E_{\text{thr}}^{\text{sph}}$. The other five rows correspond to the BH scenarios with different number of extra dimensions n and the minimal BH masses M_{\min} . The signal efficiencies ϵ_{sig} for the signal region ($N \geq 5$, $S_T \geq 7$ TeV) are shown in the last column.

method is the Residual Network (ResNet) model based on a CNN architecture. Here, the low-level detector information is converted into three-layer event images with a resolution of 50×50 bins and bin values corresponding to the energy deposits in the electromagnetic and hadronic calorimeters and p_T of tracks observed in the tracking system. We proceed to estimate the expected p -values for each hypothesis, for a given number of observed events in the signal region, originating from each of the possible scenarios.

The rest of the paper is organised as follows. In section 2 our simulation setup and event selection are explained. The six model hypotheses used in this study are also given in this section. The distributions of several observables are studied in section 3. The three ML models used in this study are laid out in section 4. Our main result, the comparison between different ML models and the exclusion p -values of the hypothesis test, is shown and discussed in section 5. Section 6 is devoted to conclusions.

2 Monte Carlo simulation and event selection

Semi-classical non-rotating BH events are generated at parton level using the **BlackMax** event generator [48]. **BlackMax** allows for the choice of the number of extra dimensions, n , and the minimal BH mass, M_{\min} . In this study, we examine five different BH hypotheses with the number of extra dimensions set to $n = 2, 4$ and 6 and the minimal BH masses to $M_{\min} = 8, 10$, and 12 TeV. The EW sphaleron/instanton-induced events are generated using **Herwig 7** [49–55] with the instanton addon library [56]. We fix the threshold energy of the sphaleron production at $E_{\text{thr}}^{\text{sph}} = 9$ TeV. We found that the sphaleron events are most similar to the BH events when the boson production is switched off. We therefore choose this option to study the most challenging scenario for discrimination. For the parton shower and hadronisation, we employ **Herwig 7** for both BH and sphaleron events. The detector response and object reconstruction are simulated with **Delphes 3** [57], assuming the conditions and geometry of the ATLAS detector. The six model hypotheses studied in this paper are listed in Table 1. In both BH and sphaleron scenarios, we treat the production cross sections as unknown parameters, since we wish to focus on the potential for discrimination, while keeping in mind, however, that the current experimental limit on the sphaleron and BH cross sections is $\sigma \lesssim 0.1$ fb [37, 47] with $E_{\text{thr}}^{\text{sph}}, M_{\min} \sim 8\text{--}12$ TeV.

Before delving into the classification problem, we would like to highlight that the SM background is under control. In particular, in Ref. [47], the CMS collaboration has performed an analysis to search for semi-classical BHs and EW sphalerons and has demonstrated that the SM background,

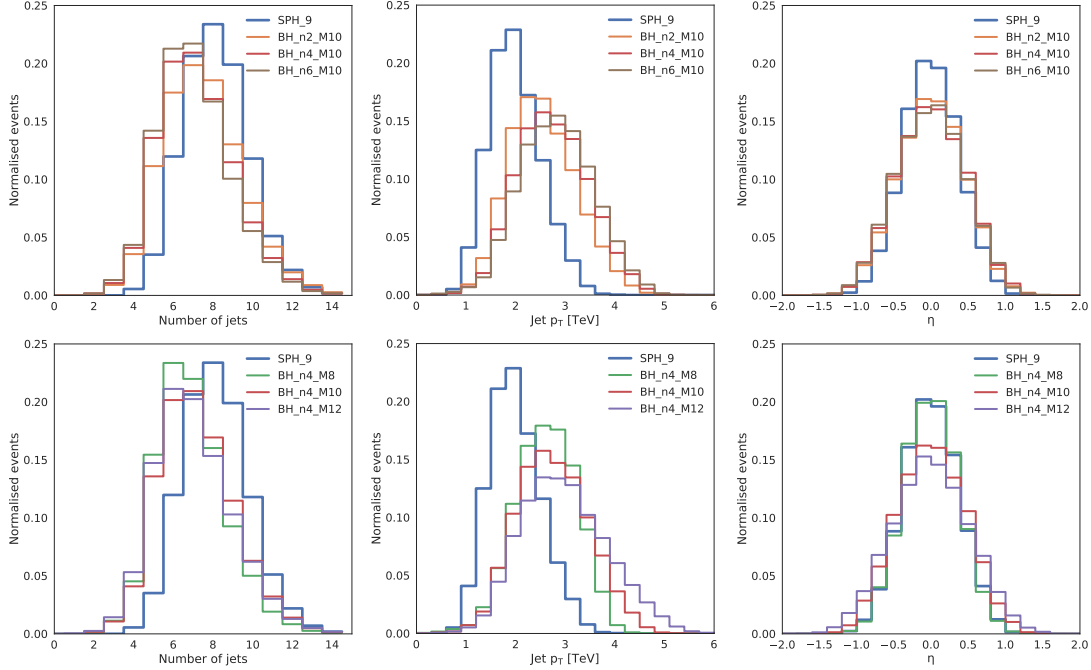


Figure 1: Normalised distributions of the number of signal jets (left), the p_T of all signal jets (middle) and the pseudorapidity η of the highest- p_T jets (right). The sphaleron events are shown with blue histograms. The plots in the upper panel show the BH scenarios with $n = 2$ (orange), 4 (red) and 6 (brown), while those in the lower panel show the BH scenarios with $M_{\min} = 8$ (green), 10 (red) and 12 (purple) TeV.

dominantly the QCD multijet events, can be heavily suppressed after an appropriate event selection. In the CMS analysis, the event selection is based on the two reconstructed quantities, N and S_T . The variable N is defined as the number of high- p_T objects in an event, with $p_T > 70$ GeV, including jets with $|\eta| < 5$, isolated electrons and photons with $|\eta| < 2.5$, and isolated muons with $|\eta| < 2.4$. In the following, we call the reconstructed objects satisfying the above criteria, *signal objects*. For example, signal electrons comprise of the isolated electrons with $p_T > 70$ GeV and $|\eta| < 2.5$. The S_T is the scalar sum of the missing transverse energy p_T^{miss} and the magnitude of transverse momenta of all N signal objects, $S_T \equiv |p_T^{\text{miss}}| + \sum_{i=1}^N |p_T^i|$. The CMS study has shown that the SM background is reduced to less than 1 event (at 13 TeV, 36 fb^{-1}) by demanding $S_T \geq 7$ TeV and $N \geq 5$. On the other hand, 17-65 % of the signal events remain after this event selection. In the following analysis, we apply essentially the same event selection. The only difference is that our signal objects are defined with a tighter pseudorapidity cut, $|\eta| < 2.4$.^{‡1} The efficiency of each model hypothesis is shown in Table 1. Since the SM background can be safely neglected by this selection, we do not consider it further in our analysis.

3 Kinematical distributions

In order to assess the sensitivity to the model hypotheses of high-level kinematical variables, we compare their distributions in Figs. 1 and 2. In the upper panels of both figures we fix the minimal BH mass at 10 TeV and vary the number of extra dimensions as $n = 2$ (orange), 4 (red) and 6 (brown), whereas in the lower panels, we fix $n = 4$ and varied $M_{\min} = 8$ (green), 10 (red) and 12 (purple) TeV.

^{‡1}We have checked the change of the signal efficiency by this modification is less than 1%.

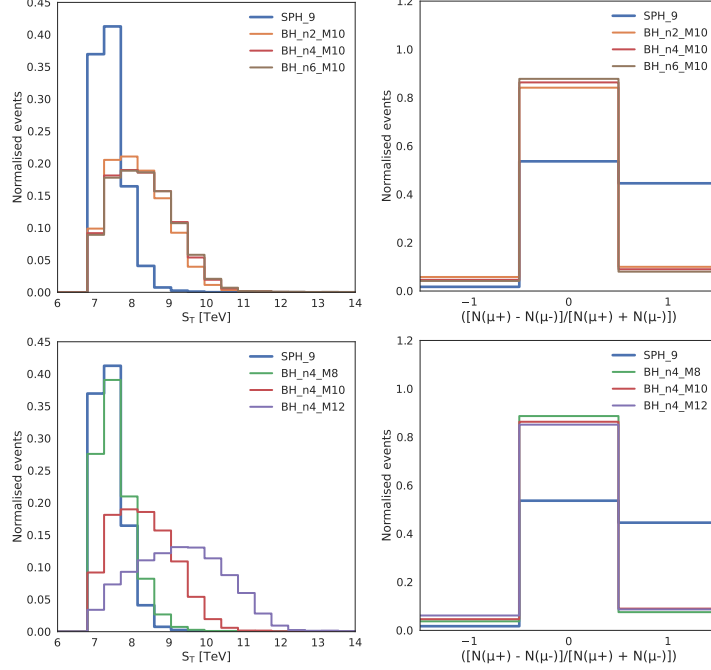


Figure 2: Normalised distributions of S_T (left) and the muon charge asymmetry CA_μ (right). The same colour coding is used as in Fig. 1.

(purple) TeV. In both panels, the distributions for sphaleron events are shown in blue.

The left panel of Fig. 1 shows the number of signal jets, where we see that the sphaleron has on average 2-3 additional signal jets than the BH scenarios. The majority ($\sim 90\%$) of sphaleron events have 6-10 signal jets and the distribution peaks at 8. On the other hand, the distributions for BH scenarios peak around 5-6 jets and have a relatively long tail towards higher multiplicities. The signal jet multiplicity is almost insensitive to the number of extra dimensions and the minimal BH mass.

The middle panel of Fig. 1 displays the accumulated p_T distributions of all signal jets. We see that the average signal jet p_T is much lower for sphaleron events. The distribution for sphaleron events has a sharper peak around 2 TeV, whereas the distributions for BH scenarios have a broader shape, and the average jet p_T is much higher, around 2.5-3.5 TeV. We observe in the upper middle plot that the jet p_T for $n = 2$ BH is slightly lower than that for $n = 4$ and 6, on average. In the lower middle plot, we also see that the jet p_T distributions have longer tails in the higher- p_T region for larger minimal BH masses.

The distributions of the pseudorapidity η of the highest- p_T jet are shown in the right panel of Fig. 1. As can be seen, the most energetic jet is produced in the central region, $|\eta| \lesssim 1$, both for sphaleron and BH events. The η distributions are similar among all scenarios, though one can observe that the sphaleron and the $(M_{\min}, n) = (8 \text{ TeV}, 4)$ BH have slightly narrower distributions, compared to the other BH scenarios.

The left panel of Fig. 2 shows the S_T distributions. As can be seen, the S_T distribution for sphalerons sharply peaks around 7.5 TeV. We observe that the S_T distribution for BH scenarios is sensitive to M_{\min} , but not to the number of extra dimensions, n . We see in the lower-left plot that the S_T distribution for the $(M_{\min}, n) = (8 \text{ TeV}, 4)$ BH is similar to that for sphalerons. The distribution gets broader and shifted towards higher energies as the minimal BH mass increases. The S_T distribution peaks around 9.5 TeV for the $(M_{\min}, n) = (12 \text{ TeV}, 4)$ BH scenario.

The right panel of Fig. 2 displays the charge asymmetry of signal muons, defined by the difference between the numbers of μ^+ and μ^- in an event, divided by the total number of muons: $CA_\mu \equiv (N_{\mu^+} - N_{\mu^-}) / (N_{\mu^+} + N_{\mu^-})$. If the event contains no signal muons, we set $CA_\mu = 0$. We see in the plots that there is a large excess of μ^+ in sphaleron events: sphaleron events have $CA_\mu = 0$ or 1 with $\sim 50\%$ chance. This asymmetry is expected because in the quark-quark initial state, the sphaleron vertex involves exactly one anti-muon or exactly one anti-muon-neutrino. Extra (anti-)muons may arise from decays of anti-top-quark or heavy hadrons. Contrary to the sphaleron case, the majority ($\sim 90\%$) of BH events has zero muon charge asymmetry. This is expected because semi-classical BH decays are essentially a thermal process. In the plots, we however observe a small excess for μ^+ all BH samples. This is a consequence of the charge conservation of BH decays and the proton contains more up-quarks rather than down-quarks especially in a large- x region of the parton distribution function.

A clear difference in several kinematical distributions between sphaleron and various BH scenarios suggests that the separation between sphaleron and BH scenarios should be relatively straightforward with sufficient collider data. Separation between different M_{\min} among BH scenarios seems also possible with some resolution, since the jet p_T and S_T distributions depend, to some extent, on M_{\min} . On the other hand, separation of the number of extra dimensions n seems challenging as no distribution in Figs. 1 and 2 exhibits a clear dependence on n . In the following sections, we investigate how well one can distinguish different scenarios, with a given number of observed events, using modern ML techniques.

4 Machine Learning setup

In this study, we examine three types of ML methods. The first two are based on the well-established XGBoost algorithm [58], which creates an ensemble of decision trees to effectively separate different events. The difference between the two is the provided input features. The other method uses the CNN based ResNet [59] architecture, which is known to be very powerful in image recognition applications.

The ML models are all trained and tested using input data created from the same datasets for training, validation and testing. The datasets consist of 10000, 3000 and 15000 events, respectively, for each physics model hypothesis, giving a total of 60000, 18000 and 90000 events. In order to assess the stability and uncertainty of classifications, for each of the three ML methods we create five independent classifiers (decision trees for XGBoost and neural networks for ResNet) by shuffling the training data at the start of each training. The mean global accuracy and its standard deviation are estimated with these five independent classifiers.

Once the training and model fine-tuning process is completed (using only the validation set to estimate the performance), the networks are used to predict the labels of the testing set. The predicted and true labels are then used to calculate the metrics for evaluation. For this study, the metric used is a simple global accuracy (number of correctly labelled events divided by the total number of events)^{‡2}.

4.1 XGBoost

The XGBoost algorithm was developed during the highly successful Higgs ML 2014 Kaggle competition (www.kaggle.com/c/higgs-boson) [61]. It is a regularizing gradient boosting framework

^{‡2}Code, results and data specifications are openly available on our project GitHub <https://github.com/choisant/imcalML> [60].

library and enables fast training of decision tree models. XGBoost is implemented in many modern HEP analyses such as the diphoton search from ATLAS [62]. As such, XGBoost is a state-of-the-art machine-learning tool for high-energy particle physics and is expected to yield good results in a short training time.

In this study, XGBoost models with two different input types have been explored. The first XGBoost method (XGBoost-High) uses the high-level inputs of the reconstructed objects. The selected input features are^{‡3} the first eight jets ($p_T^{\text{miss}}, \eta, \phi$), the first two electrons and muons (p_T, η, ϕ) and the missing transverse energy (p_T^{miss}, ϕ), which comprises 38 input features. The XGBoost algorithm tolerates missing values, but the data must be the same length for all events. The data are padded with NaN if an event does not have enough jets or leptons.

In the second XGBoost method (XGBoost-Low) low-level input is used. In this case, the input features consist of the signals in the tracking system, the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The calorimeter hits, also referred to as calorimeter towers, represent the summed-up energies at discrete angular positions across the detector. This means that the radial coordinates where the energy was deposited are not retained. The 30 highest-energy hits in each of the three different detector systems, and their corresponding angular position, were selected. This sums to 270 input features in total. Although we experimented with other variations of input features, we did not find any improvement by adding more features with lower-energy hits.

4.2 Convolutional neural network

CNNs are neural networks with convolutional layers. It is known that this architecture works extremely well in recognizing patterns in image-type data, that is data which have a matrix structure where the relative positioning of each input in the matrix is important. In this study we use the ResNet18 architecture with 18-layers.^{‡4} We convert low-level collider event data into event images as follows: starting from the 2D angular coordinates of the calorimeter towers and tracks, (ϕ, η) , we divide the range $\phi \in [-\pi, \pi], \eta \in [-5, 5]$ into 50×50 equal sized bins.^{‡5} The total energy deposit (not normalised) in the ECAL, HCAL and tracks in a given bin is assigned to their respective layer in that bin. The result is a 3-layer 2D histogram with the input shape $50 \times 50 \times 3$, which we refer to as an event image and has 7500 input features in total. In the appendix we show example event images of each of the physics scenarios in Fig. 5.

Due to the detector’s cylindrical geometry, circular padding is applied when performing convolutions. This adds pixels from the opposite edge of the image instead of the traditional zero-padding when convoluting pixels along the edge of the image.

In the course of the training process, data augmentation is applied to the images in order to artificially increase the statistics of the dataset. The augmentation consists of random rotations in the ϕ -direction and random flipping of the η axis. These transformations do not affect the physics of the events, because of the symmetries of the detector and the collision setup. Data augmentation is necessary for neural networks to be effective with the relatively small datasets used in this study. The images in the validation and testing set are not transformed in any way.

During the training, random batches of 128 images are loaded into memory at a time. Data augmentation is applied to each image, and then the network is trained using an exponential cyclic learning rate scheduler [63] with a base learning rate of 0.001, a maximum learning rate of 0.01, and $\gamma = 0.85$ and the Adam optimization algorithm [64] based on the calculated cross-entropy

^{‡3}The jets and leptons are ordered by p_T .

^{‡4}We tested models with more layers, but did not find an increase in performance.

^{‡5}Higher resolutions than 50×50 were tested. We found either no improvement or slightly worse performance.

Model	XGBoost-High	XGBoost-Low	ResNet18
Accuracy	48.91 ± 0.05	46.69 ± 0.08	50.20 ± 1.08
Training time	2 min	3.5 min	24 min
Testing time	10 s	7 s	120 s

Table 2: The global accuracy as well as the training and testing time for XGBoost-High, XGBoost-Low and ResNet18 ML methods. The mean and error of the global accuracy were estimated from an ensemble of five separately trained classifiers for each ML method. The quoted training and testing times were measured based on a GPU with an NVIDIA RTX A4500 graphic card.

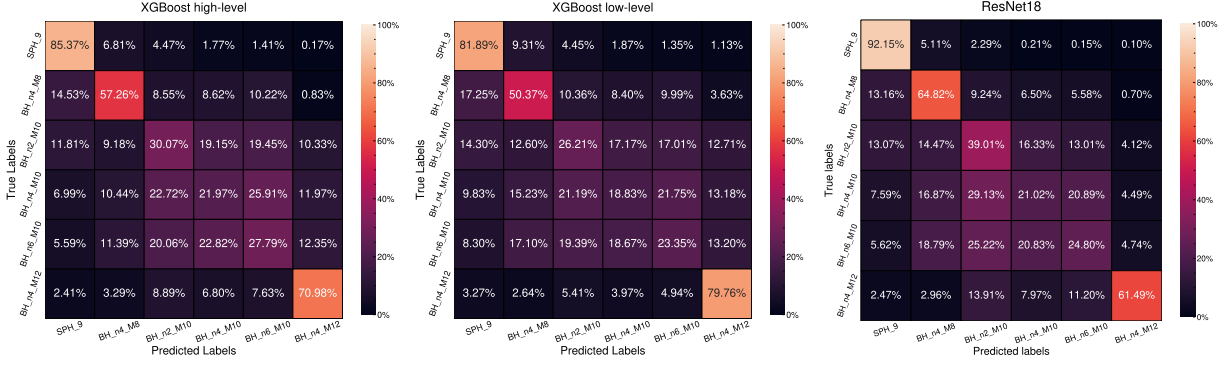


Figure 3: Confusion matrices for the XGBoost-High (left), XGBoost-Low (middle) and CNN ResNet18 (right).

loss from the batch. The network is trained for 40 epochs, below which the loss function converges around its minimum value.

5 The classification performance

The mean global accuracy with uncertainties, as well as training and testing times for each ML method, is presented in Table 2. The three ML methods we examined have individual classifier accuracy scores ranging from 46.61% (XGBoost-Low) to 51.28% (ResNet18). ResNet18 and XGBoost-High achieve similar scores and perform better than XGBoost-Low. It is interesting that ResNet18, which uses only low-level data, can perform as well as XGBoost does with high-level data inputs. It should be noted that the training and testing times for the CNN model are much longer than for XGBoost. This is a natural effect of the deep neural network's complex structure and the number of input features.

Fig. 3 shows the confusion matrices of the three ML methods we examined. The separation of sphalerons and BHs is the objective with the best performance across all models. In particular, ResNet can correctly identify the sphaleron events with $\sim 93\%$ accuracy. Furthermore, when separating the different BH samples, the models are much better at distinguishing samples with different minimal BH masses, than with different numbers of extra dimensions.

To demonstrate the practical application of the ML classifiers, we have performed a χ^2 -test for a given number of observed signal events, N_{obs} , in our signal region ($N \geq 5$, $S_T \geq 7$ TeV). Since the two best-performing ML models (ResNet and XGBoost-High) have similar classification performance, we have conducted this exercise only with the ResNet model.

For each observed event i of model hypothesis scenario J ,^{‡6} ResNet assigns a classification label, L_i , where J and L_i belong to the set of six model hypotheses in Table 1. Using the test set of 15000 events, we first create a normalised template distribution of L for each scenario J . Roughly speaking, each bin of the distribution, $P_J(L)$, is interpreted as the probability that an observed event of scenario J is classified as scenario L by the ResNet.

Having those six templates at hand and assuming that the “true” scenario is I , we performed a χ^2 test for N_{obs} observed events of scenario I (referred to as the signal sample). Specifically, we calculate the χ^2 value for the signal sample to correspond to scenario J as

$$\chi^2(I|J) = \sum_L \frac{[\mathcal{O}_I(L) - N_{\text{obs}}P_J(L)]^2}{N_{\text{obs}}P_J(L)}, \quad (5.1)$$

where $\mathcal{O}_I(L)$ is the number of events classified as L in the signal sample, $\sum_L \mathcal{O}_I(L) = N_{\text{obs}}$. The value $N_{\text{obs}}P_J(L)$ corresponds to the expected number of events labelled as L when the scenario is J . From this χ^2 , the corresponding p -value for each hypothesis can readily be calculated, with the number of degrees of freedom being [the number of hypotheses] $-1 = 5$ in this case. The p -value is interpreted as the frequentist probability that the scenario J provides the observed signal sample.

Given the current limits on the BH and EW sphaleron cross-sections [37, 47], $\sigma \lesssim 0.1 \text{ fb}$, we do not expect a large number of BH/sphalerons events can be collected in LHC Run-2, -3 and HL-LHC. For a small number of signal events, the corresponding p -values can fluctuate significantly. To understand the magnitude of the fluctuation, we performed 3000 pseudo-experiments for each N_{obs} of the “true” scenario and estimated the upper and lower errors on the p -values.

In Fig. 4, we show the result of our p -value estimation as a function of the number of observed events in the signal region. In the top-left plot, the true scenario (the origin of signal events) is SPH.9 (EW sphaleron scenario with 9 TeV threshold energy). The blue curve and the band around it show the p -value for the SPH.9 hypothesis and the standard deviation, estimated from the pseudo-experiments, respectively. As expected, the p -value is $\mathcal{O}(1)$ and do not decrease as the number of signal events increases, since it is the correct hypothesis. The other curves show the p -values for various BH hypotheses: BH_n4_M8 (green), BH_n2_M10 (orange), BH_n4_M10 (red), BH_n6_M10 (brown), BH_n4_M12 (purple). In the plot, the horizontal dashed black line corresponds to the $2\text{-}\sigma$ ($p \simeq 0.05$) exclusion and the bottom of the plot $5\text{-}\sigma$ ($p \simeq 3 \cdot 10^{-7}$). We see that just observing a few signal events is often enough to exclude all of our BH hypotheses at more than the $2\text{-}\sigma$ level. Moreover, a $5\text{-}\sigma$ exclusion can be reached if one observes 10-15 signal (SPH.9) events.

The top-middle and top-right plots show the p -values of various hypotheses when the true scenarios are BH_n4_M8 and BH_n4_M12, respectively. In both cases, we can observe that the p -values for the correct hypotheses remain $\mathcal{O}(1)$ for a larger number of signal events. We also notice that the sphaleron hypothesis and the BH hypothesis with M_{min} differ by 4 TeV from the correct hypothesis (i.e. BH_n4_M12 (purple) in the middle plot and BH_n4_M8 (green) in the right plot), are excluded very quickly as accumulating the signal events. For those scenarios, a couple of events are sufficient for a $2\text{-}\sigma$ exclusion and a $5\text{-}\sigma$ exclusion can be reached with ~ 10 events. In the same plots, we see the p -values for the BH hypotheses whose M_{min} differs by 2 TeV from the correct hypothesis. These are the scenarios with $M_{\text{min}} = 10 \text{ TeV}$ and $n = 2$ (orange), 4 (red) and 6 (brown). We see that the p -values of these three hypotheses are very similar, meaning that there is not much sensitivity to the number of extra dimensions. For those hypotheses, the p -values decrease at a somewhat slower rate: $\mathcal{O}(1)$ -20 events are required for $2\text{-}\sigma$ exclusion, while $5\text{-}\sigma$ exclusion is possible with 30-50 and 15-30 events for the true scenarios BH_n4_M8 and BH_n4_M12, respectively.

^{‡6}There is no case in which the data could contain a mix of scenarios.

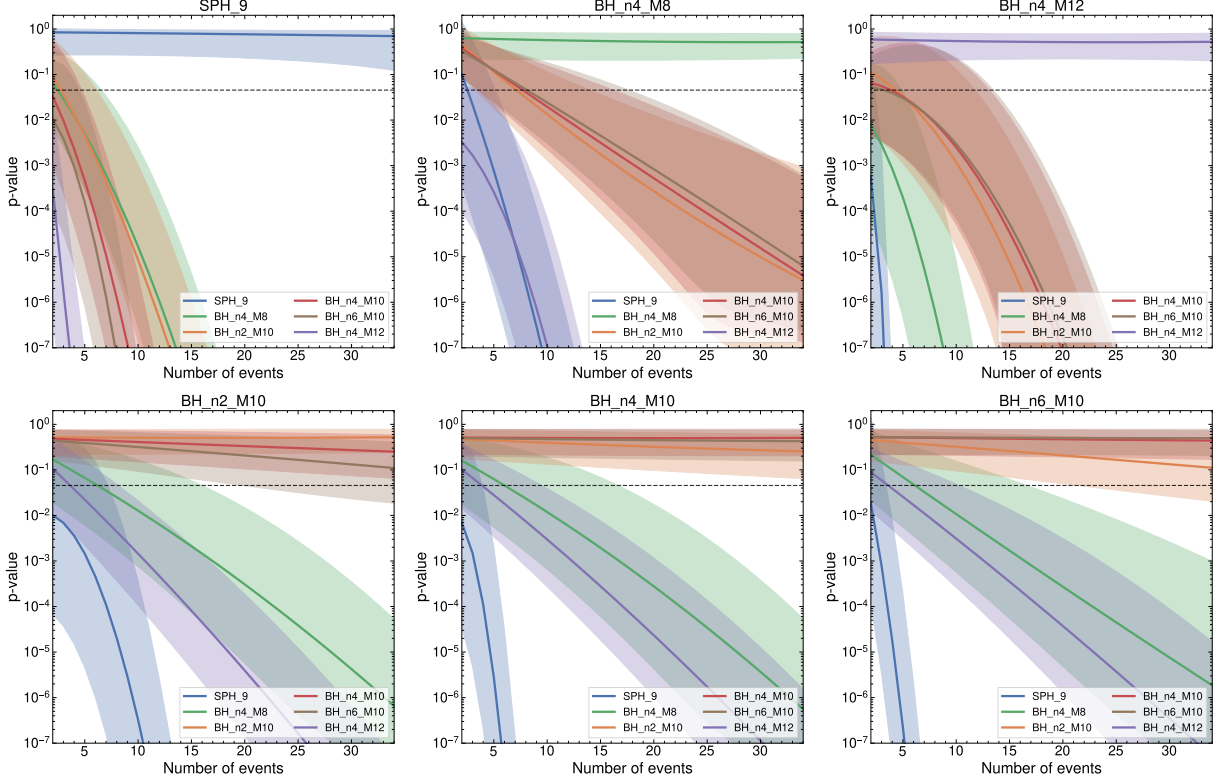


Figure 4: The exclusion p -values of six model hypotheses: SPH_9 (blue), BH_n4_M8 (green), BH_n2_M10 (orange), BH_n4_M10 (red), BH_n6_M10 (brown), BH_n4_M12 (purple) as a function of the signal events registered in the signal region. The true hypotheses are indicated at the top of each plot. The horizontal black-dashed line represents the $2\text{-}\sigma$ exclusion ($p \simeq 0.05$) and the bottom of the plots correspond to $5\text{-}\sigma$ ($p \simeq 3 \cdot 10^{-7}$).

The three plots in the bottom of Fig. 4 show the p -values when the true scenarios have $M_{\min} = 10 \text{ TeV}$ and $n = 2$ (bottom-left), 4 (bottom-middle) and 6 (bottom-right). These plots show the similar behaviour of exclusion p -values. The sphaleron hypothesis can be excluded with just a couple of signal events at the $2\text{-}\sigma$, and 6-10 events at $5\text{-}\sigma$ levels. The exclusion of the BH scenario with $M_{\min} = 12 \text{ TeV}$ (purple) is easier than that with $M_{\min} = 8 \text{ TeV}$ (green). The former requires a couple to ~ 12 events for a $2\text{-}\sigma$ exclusion, and ~ 20 -50 events to reach $5\text{-}\sigma$, while the latter scenario can be excluded with a couple to ~ 20 events at the $2\text{-}\sigma$ level, and with ~ 30 -60 events at the $5\text{-}\sigma$ levels. As can be seen from these three plots, the discrimination between the same M_{\min} and different n is more challenging. With less than 35 signal events, discrimination at the $2\text{-}\sigma$ level is less likely to be achieved. Extrapolating the slopes, we however expect that discrimination between $n = 2$ vs $n = 6$ may be possible at the $2\text{-}\sigma$ level, if more than 50 signal events are collected.

6 Conclusions

Large-multiplicity final states of jets and leptons at the LHC are ubiquitous in exotic processes, both within the Standard Model and within its extensions. In particular, final states with $\mathcal{O}(10)$ jets plus a few leptons are expected in thermal decays of semi-classical Black Holes (BHs), which are anticipated to be produced in models with large extra dimensions, characterised by two main parameters: the number of extra dimensions and the minimal mass of the BH. Similar final states are

also expected in the EW sphaleron/instanton-induced processes. In this study, we have investigated whether one can discriminate among the sphaleron scenario and five different BH scenarios listed in Table 1, through events collected at the LHC, using modern Machine Learning (ML) methods.

In studying kinematical distributions we observed that sphalerons have significantly different distributions than those of various BH scenarios, in the number of jets, p_T of jets, as well as the muon charge asymmetry. The S_T and jet p_T distributions also exhibit a dependence on the minimal BH masses. On the other hand, no clear sensitivity to the number of extra dimensions was found in any of the distributions shown in Figs. 1 and 2.

In this study, three ML models have been examined: XGBoost with low- and high-level inputs (XGBoost-Low and XGBoost-High) and the CNN based ResNet model. In the latter, the input was the low-level detector information, converted into three-layer binned event images, corresponding to the signals in the ECAL, HCAL and tracking system. We found that the XGBoost-High and ResNet models possess similar performances, though the discrimination power is the highest for the ResNet model.

To assess the practical applicability of the ML classification, we evaluated, with the ResNet model as an example, the exclusion p -value of each hypothesis, J , for a given number of observed signal events (in the signal region), originated from true scenario I . We demonstrated that the sphaleron hypothesis can be discriminated from various BH scenarios only with a few events. Separation between BH scenarios with different minimal BH masses is also possible with a reasonable number of events collected at LHC Run-2, -3 and HL-LHC. The discrimination among the BH scenarios with the same minimal BH mass but different number of extra dimensions is more challenging and requires larger number of signal events, which may be collected at future high-energy colliders.

Acknowledgements

The research of A.S.G, F.K., A.L., R.M., K.S. and T.S leading to these results has received funding from the Norwegian Financial Mechanism for years 2014-2021, grant no DEC-2019/34/H/ST2/00707. K.S. is supported by the National Science Centre, Poland, under research grant 2017/26/E/ST2/00135. R.M. is supported by the National Science Centre, Poland, under PRELUDIUM 20 research grant no 2021/41/N/ST2/00972. A.P. acknowledges support by the National Science Foundation under Grant No. PHY 2210161.

References

- [1] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “The Hierarchy problem and new dimensions at a millimeter,” *Phys. Lett. B* **429** (1998) 263–272 [[hep-ph/9803315](#)].
- [2] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “New dimensions at a millimeter to a Fermi and superstrings at a TeV,” *Phys. Lett. B* **436** (1998) 257–263 [[hep-ph/9804398](#)].
- [3] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity,” *Phys. Rev. D* **59** (1999) 086004 [[hep-ph/9807344](#)].
- [4] T. Banks and W. Fischler, “A Model for high-energy scattering in quantum gravity.” [hep-th/9906038](#).

- [5] S. Dimopoulos and G. L. Landsberg, “Black holes at the LHC,” *Phys. Rev. Lett.* **87** (2001) 161602 [[hep-ph/0106295](#)].
- [6] D. M. Eardley and S. B. Giddings, “Classical black hole production in high-energy collisions,” *Phys. Rev. D* **66** (2002) 044011 [[gr-qc/0201034](#)].
- [7] H. Yoshino and Y. Nambu, “High-energy headon collisions of particles and hoop conjecture,” *Phys. Rev. D* **66** (2002) 065004 [[gr-qc/0204060](#)].
- [8] H. Yoshino and Y. Nambu, “Black hole formation in the grazing collision of high-energy particles,” *Phys. Rev. D* **67** (2003) 024009 [[gr-qc/0209003](#)].
- [9] D. Ida, K.-y. Oda, and S. C. Park, “Rotating black holes at future colliders: Greybody factors for brane fields,” *Phys. Rev. D* **67** (2003) 064025 [[hep-th/0212108](#)]. [Erratum: *Phys. Rev. D* **69**, 049901 (2004)].
- [10] C. M. Harris, P. Richardson, and B. R. Webber, “CHARYBDIS: A Black hole event generator,” *JHEP* **08** (2003) 033 [[hep-ph/0307305](#)].
- [11] C. M. Harris, M. J. Palmer, M. A. Parker, P. Richardson, *et al.*, “Exploring higher dimensional black holes at the large hadron collider,” *JHEP* **05** (2005) 053 [[hep-ph/0411022](#)].
- [12] G. W. Gibbons and S. W. Hawking, eds., “Particle Creation by Black Holes,” *Commun. Math. Phys.* **43** (1975) 199–220. [Erratum: *Commun. Math. Phys.* **46**, 206 (1976)].
- [13] R. C. Myers and M. J. Perry, “Black Holes in Higher Dimensional Space-Times,” *Annals Phys.* **172** (1986) 304.
- [14] S. B. Giddings and S. D. Thomas, “High-energy colliders as black hole factories: The End of short distance physics,” *Phys. Rev. D* **65** (2002) 056010 [[hep-ph/0106219](#)].
- [15] B. S. DeWitt, “Quantum Field Theory in Curved Space-Time,” *Phys. Rept.* **19** (1975) 295–357.
- [16] D. N. Page, “Particle Emission Rates from a Black Hole: Massless Particles from an Uncharged, Nonrotating Hole,” *Phys. Rev. D* **13** (1976) 198–206.
- [17] N. S. Manton, “Topology in the Weinberg-Salam Theory,” *Phys. Rev. D* **28** (1983) 2019.
- [18] F. R. Klinkhamer and N. S. Manton, “A Saddle Point Solution in the Weinberg-Salam Theory,” *Phys. Rev. D* **30** (1984) 2212.
- [19] J. C. Taylor, ed., “Pseudoparticle Solutions of the Yang-Mills Equations,” *Phys. Lett. B* **59** (1975) 85–87.
- [20] M. A. Shifman, ed., “On Constrained Instantons,” *Nucl. Phys. B* **191** (1981) 429.
- [21] V. A. Rubakov and M. E. Shaposhnikov, “Electroweak baryon number nonconservation in the early universe and in high-energy collisions,” *Usp. Fiz. Nauk* **166** (1996) 493–537 [[hep-ph/9603208](#)].
- [22] A. Ringwald, “High-Energy Breakdown of Perturbation Theory in the Electroweak Instanton Sector,” *Nucl. Phys. B* **330** (1990) 1–18.

- [23] O. Espinosa, “High-Energy Behavior of Baryon and Lepton Number Violating Scattering Amplitudes and Breakdown of Unitarity in the Standard Model,” Nucl. Phys. B **343** (1990) 310–340.
- [24] P. G. Tinyakov, “Instanton like transitions in high-energy collisions,” Int. J. Mod. Phys. A **8** (1993) 1823–1886.
- [25] S. Y. Khlebnikov, V. A. Rubakov, and P. G. Tinyakov, “Instanton induced cross-sections below the sphaleron,” Nucl. Phys. B **350** (1991) 441–473.
- [26] A. H. Mueller, “On higher order semiclassical corrections to high-energy cross-sections in the one instanton sector,” Nucl. Phys. B **364** (1991) 109–126.
- [27] V. V. Khoze and A. Ringwald, “Total cross-section for anomalous fermion number violation via dispersion relation,” Nucl. Phys. B **355** (1991) 351–368.
- [28] D. Diakonov and M. V. Polyakov, “Baryon number nonconservation at high-energies and instanton interactions,” Nucl. Phys. B **389** (1993) 109–132.
- [29] I. Balitsky and A. Schafer, “Valley method versus instanton induced effective Lagrangian up to $(E/E_{\text{sphaleron}})^{8/3}$,” Nucl. Phys. B **404** (1993) 639–683 [[hep-ph/9304261](#)].
- [30] A. Ringwald, “Electroweak instantons / sphalerons at VLHC?” Phys. Lett. B **555** (2003) 227–237 [[hep-ph/0212099](#)].
- [31] F. L. Bezrukov, D. Levkov, C. Rebbi, V. A. Rubakov, and P. Tinyakov, “Semiclassical study of baryon and lepton number violation in high-energy electroweak collisions,” Phys. Rev. D **68** (2003) 036005 [[hep-ph/0304180](#)].
- [32] S. H. H. Tye and S. S. C. Wong, “Bloch Wave Function for the Periodic Sphaleron Potential and Unsuppressed Baryon and Lepton Number Violating Processes,” Phys. Rev. D **92** (2015) 045005 [[arXiv:1505.03690](#)].
- [33] J. Ellis and K. Sakurai, “Search for Sphalerons in Proton-Proton Collisions,” JHEP **04** (2016) 086 [[arXiv:1601.03654](#)].
- [34] J. Ellis, K. Sakurai, and M. Spannowsky, “Search for Sphalerons: IceCube vs. LHC,” JHEP **05** (2016) 085 [[arXiv:1603.06573](#)].
- [35] K. Funakubo, K. Fuyuto, and E. Senaha, “Does a band structure affect sphaleron processes?” [arXiv:1612.05431](#).
- [36] S. H. H. Tye and S. S. C. Wong, “Baryon Number Violating Scatterings in Laboratories,” Phys. Rev. D **96** (2017) 093004 [[arXiv:1710.07223](#)].
- [37] A. Ringwald, K. Sakurai, and B. R. Webber, “Limits on Electroweak Instanton-Induced Processes with Multiple Boson Production,” JHEP **11** (2018) 105 [[arXiv:1809.10833](#)].
- [38] D. G. Cerdeño, P. Reimitz, K. Sakurai, and C. Tamarit, “ $B + L$ violation at colliders and new physics,” JHEP **04** (2018) 076 [[arXiv:1801.03492](#)].
- [39] A. Papaefstathiou, S. Plätzer, and K. Sakurai, “On the phenomenology of sphaleron-induced processes at the LHC and beyond,” JHEP **12** (2019) 017 [[arXiv:1910.04761](#)].

- [40] M. A. Shifman, ed., “Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle,” *Phys. Rev. D* **14** (1976) 3432–3450. [Erratum: *Phys.Rev.D* 18, 2199 (1978)].
- [41] M. A. Shifman, ed., “Symmetry Breaking Through Bell-Jackiw Anomalies,” *Phys. Rev. Lett.* **37** (1976) 8–11.
- [42] HEP ML Community, “A Living Review of Machine Learning for Particle Physics.” <https://iml-wg.github.io/HEPML-LivingReview/>.
- [43] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, *et al.*, “A Convolutional Neural Network Neutrino Event Classifier,” *JINST* **11** (2016) P09001 [[arXiv:1604.01444](#)].
- [44] W. Bhimji, S. A. Farrell, T. Kurth, M. Paganini, *et al.*, “Deep Neural Networks for Physics Analysis on low-level whole-detector data at the LHC,” *J. Phys. Conf. Ser.* **1085** (2018) 042034 [[arXiv:1711.03573](#)].
- [45] M. Andrews, M. Paulini, S. Gleyzer, and B. Poczos, “Exploring End-to-end Deep Learning Applications for Event Classification at CMS,” *EPJ Web of Conferences* **214** (2019) 06031.
- [46] K. Fukushima, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biological Cybernetics* **36** (1980) 193–202.
- [47] **CMS** Collaboration, “Search for black holes and sphalerons in high-multiplicity final states in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **11** (2018) 042 [[arXiv:1805.06013](#)].
- [48] D.-C. Dai, G. Starkman, D. Stojkovic, C. Issever, *et al.*, “BlackMax: A black-hole event generator with rotation, recoil, split branes, and brane tension,” *Phys. Rev. D* **77** (2008) 076007 [[arXiv:0711.3012](#)].
- [49] M. Bahr *et al.*, “Herwig++ Physics and Manual,” *Eur. Phys. J.* **C58** (2008) 639–707 [[arXiv:0803.0883](#)].
- [50] S. Gieseke *et al.*, “Herwig++ 2.5 Release Note.” [arXiv:1102.1672](#).
- [51] K. Arnold *et al.*, “Herwig++ 2.6 Release Note.” [arXiv:1205.4902](#).
- [52] J. Bellm *et al.*, “Herwig++ 2.7 Release Note.” [arXiv:1310.6877](#).
- [53] J. Bellm *et al.*, “Herwig 7.0/Herwig++ 3.0 release note,” *Eur. Phys. J.* **C76** (2016) 196 [[arXiv:1512.01178](#)].
- [54] J. Bellm *et al.*, “Herwig 7.1 Release Note.” [arXiv:1705.06919](#).
- [55] J. Bellm *et al.*, “Herwig 7.2 release note,” *Eur. Phys. J. C* **80** (2020) 452 [[arXiv:1912.06509](#)].
- [56] A. Papaefstathiou, S. Plätzer, and K. Sakurai, “On the phenomenology of sphaleron-induced processes at the LHC and beyond,” *Journal of High Energy Physics* **2019** (2019) 17 [[arXiv:1910.04761](#)].
- [57] **DELPHES 3** Collaboration, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment,” *JHEP* **02** (2014) 057 [[arXiv:1307.6346](#)].
- [58] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’16, p. 785–794. Association for Computing Machinery, New York, NY, USA, 2016.

- [59] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” arXiv e-prints (2015) arXiv:1512.03385 [arXiv:1512.03385].
- [60] A. Grefsrud, “choisant/imcalML: 1.0,” 2023.
- [61] C. Adam-Bourdarios, G. Cowan, C. Germain-Renaud, I. Guyon, *et al.*, “The Higgs Machine Learning Challenge,” Journal of Physics: Conference Series **664** (2015) 072015.
- [62] **ATLAS** Collaboration, “Search for dark matter in events with missing transverse momentum and a Higgs boson decaying into two photons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” JHEP **10** (2021) 013 [arXiv:2104.13240].
- [63] L. N. Smith, “Cyclical Learning Rates for Training Neural Networks,” 2017.
- [64] D. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” International Conference on Learning Representations (2014) .

A Example event images

We show below in Fig. 5, event images of various physics scenarios: SPH_9, BH_n4_M8, BH_n2_M10, BH_n4_M10, BH_n6_M10 and BH_n4_M12. Three randomly-chosen in the validation dataset are used for each scenario as examples. The signals in the ECAL, HCAL and tracking system correspond to the red, green and blue colour intensities, respectively. To aid visualisation, we map the accumulated energy deposits (p_T of the tracks), E_i , in bin i to the colour intensity $I_i = f(E_i)$ with a conversion function $f(E_i) = \arctan(\ln(E_i/20 \text{ GeV}))/\pi + \frac{1}{2} \in [0, 1]$. The reconstructed jet (anti- k_T jet with radius parameter $R = 0.4$) positions for each event are indicated with orange circles of radius 0.4.

Figure 5: Event images of various physics scenarios. Three randomly-chosen events are shown for each scenario. The red, green and blue colour intensities correspond to the signals in the ECAL, HCAL and tracking system, respectively. The reconstructed jet positions for each event are indicated with orange circles of radius 0.4.

