Identifying $D_s$ decays at the LHC

E. Bakos$^1$, N. de Groot$^2$, N.Vranjes$^1$

1 Institute of Physics, University of Belgrade, Pregrevica 118, 11080, Belgrade, Serbia
2 Institute for Mathematics, Astrophysics and Particle Physics (IMAPP), Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, Netherlands and Nikhef, Netherlands

evelin.bakos@cern.ch

June 12, 2023

Abstract

In this paper we present two machine learning algorithms to identify $D_s$ mesons from radiative $W$ bosons decays at the LHC. The combined network algorithm is able to identify $D_s$ mesons via its hadronic decays with an efficiency of 67% while suppressing a background of quark and gluon jets by a factor of 100.

1 Introduction

The large amount of $W$ boson produced in $pp$ collisions at the Large Hadron Collider (LHC) enables searches for exclusive hadronic decays. These decay modes can offer precision studies of QCD factorisation [1] and are sensitive to the coupling of the $W$ boson with the photon. However, the searches for the hadronic decays are still challenging due to the large background dominated by various QCD processes. Of all these decay modes $W \rightarrow D_s \gamma$ has the largest branching fraction predicted by the Standard Model with the value of $B = (3.7 \pm 1.5) \times 10^{-8}$. No such decay has been observed so far, and the best upper limit is set by the LHCb collaboration with the value $B(W \rightarrow D_s \gamma) < 6.4 \times 10^{-4}$ at 95% confidence level [2]. The limit is obtained analysing $K^+K^\gamma\pi^+$ final states, which make up 5.4% of the $D_s$ decays. This improves on an earlier limit of $B(W \rightarrow D_s \gamma) < 1.3 \times 10^{-3}$ set by the CDF collaboration [3], using only $\phi(K^+K^-)\pi^+$ and $K^{*0}K^+$ final states, which comprise 3.9% of all $D_s$ decays. The algorithm presented in this paper offers a new approach to identify $D_s$ mesons using inclusive tagging and is sensitive to all decays at the possible expense of higher backgrounds.

A recent study [4] demonstrated that jets originating from a radiative decay of color-singlet charmonium state can be distinguished from colored jets. With machine learning algorithms we can differentiate between jets originating from radiatively produced $D_s$ mesons and background jets from gluons and quarks. With retraining, the algorithm offers an opportunity to identify other mesons originating from hadronic decays of color-singlet states as well. This would improve future searches for these rare decays and could improve the measurement precision using data to be collected during the ongoing LHC Run 3.

In the following section the simulation setup is described. Next, section 3 describes the algorithm where a Deep Neural Network (DNN), a Convolutional Neural Network (CNN), and a combined network are used to identify signal mesons. In section 4 the results are presented and an overview is given of the network performance and stability. In Section 5 prospects for the search for $W \rightarrow D_s \gamma$ are assessed.
2 Simulated samples

Proton-proton collisions are simulated at 13.6 TeV to match the Run 3 data taking period of the LHC. The sample of $D_s$ particles is obtained via the hadronic decay of the $W^\pm \rightarrow D_s\gamma$. The matrix element for the process $pp \rightarrow W$ is generated at LO accuracy in QCD using MADGRAPHv5 \cite{5,6}. The NN23LO1 PDF set \cite{7} is used in the generation. The $W \rightarrow D_s\gamma$ decay as well as parton showering and subsequent hadronisation are performed using Pythia8 \cite{8} with the A14 ATLAS tune \cite{9}.

The main background processes (in terms of $D_s$ identification) are $pp \rightarrow gg$ and $pp \rightarrow qq$ where $g$ and $q$ denotes a gluon and a quark respectively. The background samples are modelled separately using the same setup as for the signal events.

The detector response is simulated via the Delphes \cite{10} package using the ATLAS detector configuration files. Jets are reconstructed as pFlow jets with the anti-$k_t$ \cite{11} jet clustering algorithm with $\Delta R = 0.4$ and are required to satisfy $p_T > 25$ GeV and $|\eta| < 2.1$ selection criteria. Jets are considered as a $D_s$ meson if the angular distance to the truth $D_s$ particle is $\Delta R < 0.2$. The entire configuration can be found in \cite{12}.

3 $D_s$ identification using machine learning algorithm

The full set of $W \rightarrow D_s\gamma$ signal sample consists of 160k events, while the $qq$ and the $gg$ background sample both contains 80k events. This makes the full background sample with 160k events comparable to the signal. Before the training all the samples were divided into training and testing set, consisting of 70% and 30% of the full dataset respectively. To create the machine learning algorithm TensorFlow \cite{13} and Keras \cite{14} libraries were used. To determine the model performance we use the receiver operating characteristic (ROC) curve and in particular the area under the ROC curve (AuC). The network hyper-parameters were optimized with grid search to make sure that the best performing models are used to obtain the results.

3.1 Deep Neural Network

Signal jets originate from the decay of an isolated $D_s$ (not surrounded by fragmentation tracks) and will be more collimated than background jets. This is particularly true for gluon jets, since gluon-initiated jets have higher particle multiplicity and a softer fragmentation function, due to the large color factor. Variables $\Delta \phi$ and $\Delta \eta$, which measure the width of the jet in the $\phi$ and $\eta$ direction, as well as $R_{em}$ and $R_{track}$ which measure the $\Delta R$ with respect to the jet axis in case of tracks and electromagnetic clusters can be used to distinguish jets originating from $D_s$ from the background jets. Multiplicity of charged and neutral particles ($n_{ch}$ and $n_0$) in jets originating from $D_s$ is lower compared to jets from quarks and gluons. From the lower constituent multiplicity it can also be deducted that signal jets have lower invariant mass. The $m_{tr}$ measures the invariant mass of all charged tracks while $m_{ij}$ defines the invariant mass of all constituents in the jet. Jets emerging from $D_s$ mesons are also less surrounded with hadronic activity caused by the fragmentation. The $p_{core}$ and $f_{core}$ measure the ratio of sum $p_T$ in a cone and the jet $p_T$ and the ratio of sum $E_T$ in a cone and the jet total $E_T$ respectively. The $E_{had}/E_{em}$ defines the energy ratio in the hadronic and the electromagnetic calorimeter.

We start with the variables used in \cite{4}. These variables are further extended with the absolute values of the total charge and the jet-charge ($p_T$ weighted charge sum \cite{15}). The charge is expected to peak at zero for gluon jets, at one for signal jets and have
a higher average value for quark jets. In addition, with the b-jet tagging we gain some discriminating power against b-jets. A particular class of generalized angularities ($\lambda_k^\beta$) \[16\] are also added to the algorithm, which are efficient in distinguishing quark jets from gluon jets.

Furthermore, the N-Subjettiness \[17\] is also used, which measures to what degree the jet is composed of N subjets. For our signal jets the N-subjettiness expected to be close to zero, since all the radiation is aligned with the direction of the jet, meaning N (or fewer) subjets. $gg$ background jets have $\tau_N >> 0$, since large fraction of their energy is distributed away from the jet direction. All the variables used for the ML algorithm are listed in Table 1 and also shown in Figure 1.

| Name     | Description                                                                 |
|----------|-----------------------------------------------------------------------------|
| $\Delta \eta$ | width of the jet in $\eta$                                                   |
| $\Delta \phi$ | width of the jet in $\phi$                                                   |
| $m_{tr}$  | invariant mass of all charged tracks in the jet                            |
| $m_j$     | invariant mass of all constituents of the jet                               |
| $n_{ch}$  | charged particle multiplicity                                                |
| $n_0$     | neutral particle multiplicity                                                |
| $|Q|$     | absolute value of the total charge                                          |
| $|q_j|$        | jet charge ($p_T$ weighted charge sum, $\Sigma_i q_i \cdot p_{T_i}/\Sigma_i p_{T_i}^{1/2}$) |
| b-tag     | output of the b-tagging algorithm                                            |
| $R_{em}$  | average $\Delta R$ with respect to the jet axis weighted by electromagnetic energy |
| $R_{track}$ | $p_T$ weighted average $\Delta R$ for tracks                              |
| $f_{em}$  | fraction of EM energy over total neutral energy of the jet                  |
| $p_{core1}$ | ratio of sum $p_T$ in a cone of $\Delta R < 0.1$ and the jet $p_T$         |
| $p_{core2}$ | ratio of sum $p_T$ in a cone of $\Delta R < 0.2$ and the jet $p_T$         |
| $f_{core1}$ | ratio of sum ET in a cone of $\Delta R < 0.1$ and the jet total ET          |
| $f_{core2}$ | ratio of sum ET in a cone of $\Delta R < 0.2$ and the jet total ET          |
| $f_{core3}$ | ratio of sum ET in a cone of $\Delta R < 0.3$ and the jet total ET          |
| $(p_{T}^2)^2$ | $\lambda_0^2$                                                             |
| LHA       | Les Houches Angularity; $\lambda_{0.5}^1$                                 |
| Width     | $\lambda_1^1$                                                             |
| Mass      | $\lambda_2^1$                                                             |
| $E_{had}/E_{em}$ | ratio of the hadronic versus electromagnetic energy deposited in the calorimeter |
| $\tau_0$, $\tau_1$, $\tau_2$ | N-Subjettiness                                                              |

Table 1: DNN input variables.

Based on the optimisation results, the final model consists of one input layer and two hidden layers with 35, 20 and 12 nodes respectively. The activation function for the input layer and both hidden layer is $\tanh$. As is common with classification problems, the output layer is activated with the $\text{sigmoid}$ function. The full set of hyperparameters is summarized in Table 2. A feature importance plot for the DNN network is also presented in Figure 2. It can be seen that the most important features are the jet constituents invariant mass and the N-Subjettiness, while the b-tag variable has no impact on the network at all. This indicates kinematics of the generated sample does not major impact on the obtained results.

### 3.2 Convolutional Neural Network

Another approach for developing a $D_s$ identification algorithm is to use a CNN. In this case the input variables are low level variables: energy deposit in the electromagnetic and the hadronic calorimeter and track transverse momentum, which are plotted as a 3D image. The advantage of this approach is that one can use relatively raw data instead of
Figure 1: Distributions of the variables used for $D_s$ identification, using DNN. The signal is presented with a solid blue line, while the $gg$ and $qq$ backgrounds are drawn with dashed orange and dotted green lines respectively.
Figure 2: Feature importance plot of DNN. The blue bars represent the weight of each feature (variable) within the network.

carefully constructed variables.

In the context of this analysis, these energy deposits and the track transverse momentum are converted into a \( 20 \times 20 \) jet image. Since the jet reconstruction parameter is \( \Delta R = 0.4 \), and the segmentation of the ECAL is \( 0.02 \times 0.02 \), the grid size of the jet image is equal to the smallest possible tower size in the \( \eta-\phi \) plane. The variables introduced in three different channels as it is the case of an RGB picture, where the hadronic deposit is noted with blue, the electromagnetic deposit with green and the track transverse momentum with red. The schematic illustration of the jet image is shown in Figure 3.

![Jet Image Construction](image)

(a) Signal          (b) Background

Figure 3: Jet image construction from low level variables. The hadronic deposit is noted with blue back slash pattern, the electromagnetic deposit with dotted green and the track transverse momentum with red with forward slash pattern composing an RGB input picture to the CNN algorithm.

Our CNN model consists of 5 layers: 3 convolutional and 2 fully connected dense layer. The number of nodes in the convolutional layers are 30, 8 and 8 respectively. The window
sizes are \([3 \times 3]\) and \([5 \times 5]\) in the last layer, while the activation function is \text{tanh} in all three cases. A maxpooling layer is added after the second convolutional layer. The number of nodes in the first dense layers is 10 with the \text{ReLU} activation function. The output layer is again a dense \text{sigmoid}. The parameters of the final CNN model are summarized in Table 2.

### 3.3 Combined network

To further improve the efficiency of the network, the DNN and the CNN models are merged into a single network. In this case, the output of the DNN and the output of the CNN are the inputs of the next combined layer. The last layer of the model performs the classification and the results depend both on the output of the CNN and the DNN.

The best performing combined network has slightly different number of nodes within the DNN layers: 33, 20 and 14 respectively. Another significant change compared to the previously introduced models is the absence of the dense layers after the convolutional layers. Instead a combined dense layer is introduced with 8 nodes and \text{ReLU} activation function. The classification happens in the last \text{sigmoid} layer. The parameters of the combined model are summarized in the last column of Table 2.

| Parameter                        | DNN                  | CNN                  | Combined              |
|----------------------------------|----------------------|----------------------|-----------------------|
| Dense layer nodes                | 35 - 20 - 12 - 1     | –                    | 33 - 20 - 14          |
| Dense layer activation           | tanh - tanh - tanh - sigmoid | –                    | tanh - tanh - tanh |
| Convolutional layer nodes        | 30 - 8 - 8           | [3\times3], [3\times3], [5\times5] | 30 - 8 - 8           |
| Window size                      | –                    | \[3\times3], [3\times3], [5\times5] | \[3\times3], [3\times3], [5\times5] |
| Convolutional layer activation   | –                    | tanh - tanh - tanh   | tanh - tanh - tanh   |
| Max pooling                      | –                    | After the 1\text{st} convolutional layer | –                    |
| Dense layers after convolution    | –                    | 10(relu) - 1(sigmoid) | –                    |
| Combined layer nodes             | –                    | –                    | 8 - 1                 |
| Combined layer activation        | –                    | –                    | relu - sigmoid        |
| Loss function                    | binary cross-entropy | Adam                | 40                   |
| Optimizer                        |                       |                      | Batch size            |
| Training epochs                  |                      | 40                   | Batch size            |

Table 2: Hyperparameters of the different network types.

### 4 Results

The ROC curves of the different models are presented in Figure 4, while the output distributions of the models can be seen in Figure 5. Table 3 shows the AuC values of the different networks defined previously. As is expected, the combined model performs the best with 0.980, which corresponds to a signal efficiency of 67(25)% at a background rejection factor of 100(1000). Using DNN only one can reach a signal efficiency of 52(21)%, while using only CNN the efficiency is 52(14)% at 100(1000) times background rejection. As it can be seen, the performance is significantly better against a single background of gluon jets then against quark jets. This can be further improved if one uses only a gluon sample for training to an AuC of 0.994.

We examined the response of the network to various samples not used in the training. First we looked at a charm quark sample to see if the \(D_s\) events would be tagged as signal at a higher rate. For a network cut-off of 0.5 we find a positive tag for signal events of 94%. For of the generic charm jets this was 15% and 17% when a truth \(D_s\) was matched to jet. This result is not materially different from the generic quark-jet sample with 17% and indicates that the absence of fragmentation tracks around the jets and a narrow jet...
With low multiplicity are more important than the exact D-meson decay topology. Bottom-quark jets are tagged at a significantly lower rate: 6% in general and 2% when a $D_s$ meson was present. For hadronic $\tau$ decays we expect a high tagging rate, since they are quite similar to our signal events. Indeed, we find that 64% of the hadronic $\tau$ jets are tagged as signal. This is not surprising, given that $\tau$ leptons are also produced in a color-singlet state and more than 5% of the $D_s$ mesons decay to $\tau$s.

We also investigate the stability of the network performance under variations of the simulation parameters. To study this, we apply the recommended variations of the Pythia8 framework. These variations cover a range of data observable: variation 1 is related to the underlying event activity, variation 2 is covering the jet shapes and substructure and the three variations 3 cover the effects of initial and final state radiation. The results of the variance in the model performance is presented in Table 4.

![ROC curves for the different network types.](image)

**Figure 4:** ROC curves for the different network types.

Table 3: Overview of the training results using the combined network. Mixed background test samples contain 50% quark and 50% gluon jets.

| Network type | Test sample   | Training sample | AuC  |
|--------------|---------------|-----------------|------|
| DNN          | $D_s$ vs mixed| $D_s$ vs mixed  | 0.969|
| CNN          | $D_s$ vs mixed| $D_s$ vs mixed  | 0.974|
|              | $D_s$ vs gluon| $D_s$ vs gluon  | 0.994|
| Combined     | $D_s$ vs quark| $D_s$ vs mixed  | 0.964|
|              | $D_s$ vs gluon| $D_s$ vs gluon  | 0.994|
|              | $D_s$ vs quark| $D_s$ vs quark  | 0.965|

![Output of the different networks for signal (blue) and background (red).](image)

**Figure 5:** Output of the different networks for signal (blue) and background (red).
Parameter & +variation & -variation \\
\hline
Var1: UE activity & -0.008 & 0.003 \\
Var2: jet shapes and substructure & -0.001 & 0.010 \\
Var3a: ISR/FSR $t\bar{t}$ gap & -0.002 & 0.007 \\
Var3b: ISR/FSR 3/2 jet ratio & -0.011 & 0.002 \\
Var3c: ISR & -0.007 & 0.006 \\
\hline
Table 4: Variations in the AuC for different Pythia8 tunes.

The effect of pile-up is also taken account during the analysis. Within the Delphes framework the additional tracking and vertexing information is not available, meaning that our estimate is worse than the real life conditions on the LHC experiments. We simulated samples with pile-up of $\langle \mu \rangle = 40$ meaning on average 40 pile-up interaction, which is the expected amount for LHC Run 3 conditions. The retrained network, without further optimisation shows a drop of 0.076 in the AuC, meaning that while pile-up has a significant effect, the model is still able to identify $D_s$ mesons. One can note however, that pileup mitigations techniques implemented in Delphes are suboptimal, hence the expected effect with real data is smaller.

5 Prospects

In this section prospects for the measurement of $\mathcal{B}(W \to D_s\gamma)$ using the method described previously is studied. For the purpose of this exercise it is assumed that low-pileup data corresponding to the integrated luminosity of 1 fb$^{-1}$ is collected during LHC Run 3. Events are required to have one jet tagged as $D_s$ and an isolated photon with $p_T > 30$ GeV. Events with invariant mass of jet-photon system $\pm 10$ GeV around $W$ boson mass are selected. Triggering efficiency is assumed to be 100%. Total signal efficiency for $W^+ \to D_s\gamma$ ($W^- \to D_s\gamma$) is estimated to be 16.8% (20.3%) respectively.

In order to estimate background level, large MC samples of $pp \to gg$ and $pp \to qq$, as well as $pp \to q\bar{q}$, $Z \to ee$ and $Z \to \tau\tau$ are generated with MADGRAPHv5 and Pythia8. The detector response is simulated via Delphes package using the ATLAS detector configuration files. Backgrounds are normalised according to their generated cross sections. The total level of background is estimated to be 1.39M events corresponding to the integrated luminosity of 1 fb$^{-1}$. The background is dominated by QCD process while only 2% the total background arises from $Z$ boson events.

The $CL_s$ method [18][19] is used to calculate upper limit on the branching fraction of the $W \to D_s\gamma$ decay. Signal uncertainty is assumed to be 10% and has only marginal impact on the calculated limit. The uncertainty on the background level is assumed to be 0.5% as obtained in the ATLAS search for radiative Higgs boson decay [20]. The calculated $CL_s$ exclusion as a function of branching fraction of $W \to D_s\gamma$ is shown in Figure 6.

The expected upper limit at the 95% confidence level is determined to be:

$$\mathcal{B}(W \to D_s\gamma) < 4.24^{+0.25}_{-0.28} \times 10^{-4},$$

which is comparable to the observed upper limit from LHCb.

With the entire Run 3 dataset corresponding to about 300 fb$^{-1}$, assuming trigger efficiency of 40% and taking into account deterioration of the $D_s$ tagger due to high pileup the expected upper limit improves to $\mathcal{B}(W \to D_s\gamma) < 1.2 \times 10^{-4}$. Development of dedicated trigger is needed to achieve corresponding precision.
Figure 6: Expected upper limit on branching fraction of the $W \to D_s\gamma$ decay.

6 Conclusion

The algorithm to identify jets originating from $D_s$ mesons in radiative $W$ decays presented in this paper shows a good efficiency of 67% for signal with a 100 times rejection of a jets from quarks and gluons. Against a single background of gluon jets the algorithm works even better. The algorithm is stable under the variations of the simulation parameters and it also works in the presence of pile-up but at a significant loss of performance. The algorithm opens up the possibility to further improve measurements and searches involving $D_s$ mesons, especially in case of the rare decays that suffer from low statistics. We find very similar performance for a deep neural network and a convolitional neural network. The combined network performs slightly better than either. With low pileup dataset corresponding to the integrated luminosity of 1 fb$^{-1}$ upper limit on branching fraction of $W \to D_s\gamma$ decay can be determined at the level of $\mathcal{B}(W \to D_s\gamma) < 4.2 \times 10^{-4}$.

Acknowledgments

This work was partially supported by the Serbian Ministry of Education, Science and Technological development and the EU Erasmus plus programme.

References

[1] Y. Grossman, M. König and M. Neubert, *Exclusive Radiative Decays of W and Z Bosons in QCD Factorization*, JHEP 04, 101 (2015), doi:10.1007/JHEP04(2015)101, arXiv:1501.06569.

[2] *Search for the rare decays $W^+ \to D_s^+\gamma$ and $Z \to D^0\gamma$ at LHCb* (2022), arXiv:2212.07120.

[3] F. Abe et al., *Search for the rare decay $W^\pm \to D_s^{\pm}\gamma$ in pp collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D 58, 091101 (1998), doi:10.1103/PhysRevD.58.091101.

[4] N. de Groot and S. Castells, *Identifying hadronic charmonium decays in hadron colliders*, SciPost Phys. Core 2, 8 (2020), doi:10.21468/SciPostPhysCore.2.2.008.
[5] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater and T. Stelzer, MadGraph/MadEvent v4: The New Web Generation, JHEP 09, 028 (2007), doi:10.1088/1126-6708/2007/09/028 [arXiv:0706.2334]

[6] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, JHEP 07, 079 (2014), doi:10.1007/JHEP07(2014)079 [arXiv:1405.0301]

[7] R. D. Ball et al., Parton distributions from high-precision collider data, Eur. Phys. J. C 77(10), 663 (2017), doi:10.1140/epjc/s10052-017-5199-5 [arXiv:1706.00428]

[8] T. Sjostrand, S. Mrenna and P. Z. Skands, A Brief Introduction to PYTHIA 8.1, Comput. Phys. Commun. 178, 852 (2008), doi:10.1016/j.cpc.2008.01.036 [arXiv:0710.3820].

[9] ATLAS Pythia 8 tunes to 7 TeV data, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva (2014).

[10] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens and M. Selvaggi, DELPHES 3, A modular framework for fast simulation of a generic collider experiment, JHEP 02, 057 (2014), doi:10.1007/JHEP02(2014)057 [arXiv:1307.6346].

[11] M. Cacciari, G. P. Salam and G. Soyez, The anti-\( k_t \) jet clustering algorithm, JHEP 04, 063 (2008), doi:10.1088/1126-6708/2008/04/063 [arXiv:0802.1189].

[12] E. Bakos, DsGammaAnalysis, https://github.com/ebakos/DsGammaAnalysis (2022).

[13] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg et al., Tensorflow: A system for large-scale machine learning, CoRR abs/1605.08695 (2016), arXiv:1605.08695.

[14] F. Chollet et al., Keras, https://github.com/fchollet/keras (2015).

[15] R. D. Field and R. P. Feynman, A Parametrization of the Properties of Quark Jets, Nucl. Phys. B 136, 1 (1978), doi:10.1016/0550-3213(78)90015-9.

[16] A. J. Larkoski, J. Thaler and W. J. Waalewijn, Gaining (Mutual) Information about Quark/Gluon Discrimination, JHEP 11, 129 (2014), doi:10.1007/JHEP11(2014)129 [arXiv:1408.3122].

[17] J. Thaler and K. Van Tilburg, Identifying Boosted Objects with N-subjettiness, JHEP 03, 015 (2011), doi:10.1007/JHEP03(2011)015 [arXiv:1011.2268].

[18] A. L. Read, Presentation of search results: The CL\(_s\) technique, J. Phys. G 28, 2693 (2002), doi:10.1088/0954-3899/28/10/313.

[19] T. Junk, Confidence level computation for combining searches with small statistics, Nucl. Instrum. Meth. A 434, 435 (1999), doi:10.1016/S0168-9002(99)00498-2 [arXiv:hep-ex/9902006].

[20] M. Aaboud et al., Search for exclusive Higgs and Z boson decays to \( \phi\gamma \) and \( \rho\gamma \) with the ATLAS detector, JHEP 07, 127 (2018), doi:10.1007/JHEP07(2018)127 [arXiv:1712.02758].

10