

RECEIVED: December 24, 2018

REVISED: February 7, 2019

ACCEPTED: February 14, 2019

PUBLISHED: February 20, 2019

# Searching for dark matter in final states with two jets and missing transverse energy

Ulrich Haisch<sup>a</sup> and Giacomo Polesello<sup>b</sup>

<sup>a</sup>Max Planck Institute for Physics,  
Föhringer Ring 6, 80805 München, Germany

<sup>b</sup>INFN — Sezione di Pavia,  
Via Bassi 6, 27100 Pavia, Italy

E-mail: [haisch@mpp.mpg.de](mailto:haisch@mpp.mpg.de), [giacomo.polesello@cern.ch](mailto:giacomo.polesello@cern.ch)

**ABSTRACT:** We reemphasise the usefulness of angular correlations in LHC searches for missing transverse energy ( $E_T^{\text{miss}}$ ) signatures that involve jet ( $j$ ) pairs with large invariant mass. For the case of mono-jet production via gluon-fusion, we develop a realistic analysis strategy that allows to split the dark matter (DM) signal into distinct one jet-like and two jet-like event samples. By performing state-of-the-art Monte Carlo simulations of both the mono-jet signature and the standard model background, it is shown that the dijet azimuthal angle difference  $\Delta\phi_{j_1j_2}$  in  $2j + E_T^{\text{miss}}$  production provides a powerful discriminant in realistic searches. Employing a shape fit to  $\Delta\phi_{j_1j_2}$ , we then determine the LHC reach of the mono-jet channel in the context of spin-0  $s$ -channel DM simplified models. The constraints obtained by the proposed  $\Delta\phi_{j_1j_2}$  shape fit turn out to be significantly more stringent than those that derive from standard  $E_T^{\text{miss}}$  shape analyses.

**KEYWORDS:** Dark matter, Hadron-Hadron scattering (experiments), Jets

ARXIV EPRINT: [1812.08129](https://arxiv.org/abs/1812.08129)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>DM signal and SM background</b>	<b>2</b>
<b>3</b>	<b>Analysis strategy</b>	<b>4</b>
<b>4</b>	<b>LHC Run-3 and HL-LHC projections</b>	<b>6</b>
<b>5</b>	<b>Conclusions and outlook</b>	<b>7</b>
<b>A</b>	<b>Supplementary material</b>	<b>9</b>

---

**1 Introduction**

One of the main channels used in the search for direct production of dark matter (DM) at hadron colliders is the final state that includes a high transverse momentum ( $p_T$ ) hardonic jet recoiling against the undetectable DM particles. These so-called mono-jet searches have a long and partly colourful history. They have been performed in the past at all major general purpose detector experiments like UA1, CDF and DØ [1–3], and at the LHC they now represent an important pillar of the search strategy for new physics beyond the standard model (SM).

The latest ATLAS and CMS mono-jet results [4, 5] are based on an integrated luminosity of around  $36 \text{ fb}^{-1}$  of 13 TeV LHC data. Both searches require the presence of at least a single high- $p_T$  jet and fit the shape of the missing transverse energy ( $E_T^{\text{miss}}$ ) spectrum to extract limits on DM production. Given the good theoretical understanding of the  $Z/W$ +jets SM backgrounds [6], LHC measurements of the  $E_T^{\text{miss}}$  distribution in mono-jet production place the leading constraints on  $s$ -channel DM simplified models (see for instance [7–9] and references therein) in certain regions of parameter space. However, since the corresponding shapes of the  $E_T^{\text{miss}}$  spectra are featureless and largely independent of the type of mediation mechanism, existing mono-jet measurements provide insufficient information to determine additional DM properties.

In order to disentangle different types of DM-SM interactions through studies of  $X + E_T^{\text{miss}}$  signatures more complicated observables and/or final states  $X$  need to be considered. The simplest option are channels with at least two SM particles such as two jets ( $j$ ) or two charged leptons ( $l$ ) and  $E_T^{\text{miss}}$  in the final state. The  $2j + E_T^{\text{miss}}$  signature can thereby result from either the gluon-fusion [10] or the vector-boson fusion [11–13] channel, while the  $2l + E_T^{\text{miss}}$  signal can for instance arise from  $t\bar{t} + E_T^{\text{miss}}$  [14, 15] and  $tW + E_T^{\text{miss}}$  [16, 17] production. In all cases the angular correlations of the visible final state particles have

been studied and shown to provide useful information on the structure of the interactions between the dark and the SM sector. The goal of this work is to reassess the usefulness of dijet angular correlations in the gluon-fusion  $2j + E_T^{\text{miss}}$  channel, applying the general ideas presented in [10] to the case of spin-0 simplified DM models. By performing simulations of the DM signal taking into account the effects of matrix element matching and merging, parton shower and hadronisation corrections and a realistic detector modelling, we show that the azimuthal angle difference  $\Delta\phi_{j_1j_2}$  between the two jets in  $2j + E_T^{\text{miss}}$  events furnishes a powerful model discriminant in a realistic experimental environment. We find that compared to standard  $E_T^{\text{miss}}$  likelihood fits, the inclusion of shape information on  $\Delta\phi_{j_1j_2}$  should lead to a significantly improved reach in spin-0  $s$ -channel DM simplified models. Projections are presented based on  $300 \text{ fb}^{-1}$  and  $3 \text{ ab}^{-1}$  of 14 TeV LHC data, corresponding to LHC Run-3 and the high-luminosity phase of the LHC (HL-LHC).

Our article is organised as follows. In section 2 we briefly describe the structure of the DM simplified models that we use to interpret the  $2j + E_T^{\text{miss}}$  searches. In this section also the generation of the DM signal and the SM backgrounds is explained and our detector simulation is described. Section 3 details our analysis strategy and discusses the angular correlations of the DM signal. The LHC Run-3 and HL-LHC projections are presented in section 4. We conclude in section 5. Supplementary material can be found in appendix A.

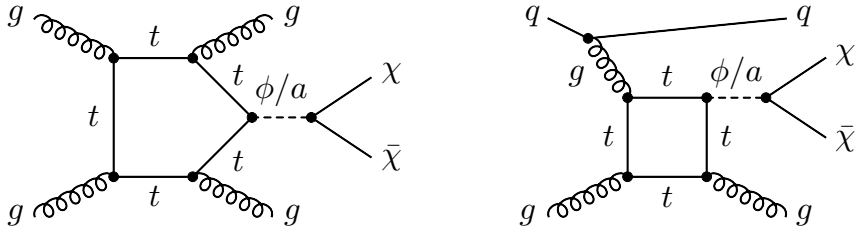
## 2 DM signal and SM background

In our work the following simplified Lagrangians are studied (see e.g. [7–9])

$$\begin{aligned} \mathcal{L}_\phi &\supset -g_\chi \phi \bar{\chi} \chi - \frac{\phi}{\sqrt{2}} \sum_{q=u,d,s,c,b,t} g_q y_q \bar{q} q, \\ \mathcal{L}_a &\supset -i g_\chi a \bar{\chi} \gamma_5 \chi - i \frac{a}{\sqrt{2}} \sum_{q=u,d,s,c,b,t} g_q y_q \bar{q} \gamma_5 q, \end{aligned} \tag{2.1}$$

which describe the coupling of a dark sector to the SM through the  $s$ -channel exchange of scalar ( $\phi$ ) and pseudoscalar ( $a$ ) mediators. In (2.1) the symbol  $\chi$  represents the DM particle assumed to be a Dirac fermion,  $g_\chi$  is a dark-sector Yukawa coupling,  $y_q = \sqrt{2}m_q/v$  are the SM quark Yukawa couplings with  $m_q$  the mass of the relevant quark  $q$  and  $v \simeq 246 \text{ GeV}$  the Higgs vacuum expectation value, and  $\gamma_5$  denotes the fifth Dirac matrix.

The DM signal samples are generated at leading order (LO) using the `DMsimp` [18] implementation of the Lagrangians (2.1) together with `MadGraph5_aMC@NLO` [19] and `NNPDF3.0` [20] parton distribution functions (PDFs). The associated production of DM with one and two jets are generated for  $pp$  collisions at a centre-of-mass (CM) energy of 14 TeV. See figure 1 for representative one-loop graphs that contribute to the  $2j + E_T^{\text{miss}}$  signal in the spin-0  $s$ -channel DM simplified models. The events are showered with `PYTHIA 8.2` [21] using the Catani-Kraus-Kuhn-Webber (CKKW) jet matching prescription [22]. We consider five different values of the mediator mass  $M_{\phi/a}$  in the range from 50 GeV to 1 TeV. The mass of the DM particles is set to  $m_\chi = 1 \text{ GeV}$  and we employ  $g_\chi = g_t = 1$  for the couplings of the mediators to DM and top quarks. The total decay width  $\Gamma_{\phi/a}$  of the mediator



**Figure 1.** Two examples of Feynman diagrams describing  $2j + E_T^{\text{miss}}$  production in the spin-0  $s$ -channel DM simplified models (2.1).

is assumed to be minimal and calculated at tree level using `MadGraph5_aMC@NLO`. Since in the narrow width approximation the signal predictions factorise into the cross sections for  $pp \rightarrow 2j + \phi/a$  production times the  $\phi/a \rightarrow \chi\bar{\chi}$  branching ratio, changing  $\Gamma_{\phi/a}$  leads only to a rescaling of the signal strength. The experimental acceptance is instead insensitive to the total decay width, and hence it is sufficient to generate samples for a single choice of couplings. The predictions for other values of  $g_\chi$  and  $g_t$  can then be obtained by scaling with the associated  $\phi/a \rightarrow \chi\bar{\chi}$  branching ratio.

The dominant SM backgrounds arise from  $Z/W + \text{jets}$  production. We consider separately the  $Z + \text{jets}$  channel with  $Z \rightarrow \nu\bar{\nu}$  and the  $W + \text{jets}$  mode with  $W \rightarrow l\nu$  where  $l = e, \mu, \tau$ . These backgrounds are generated at LO with `MadGraph5_aMC@NLO` and `NNPDF3.0` PDFs, and can contain up to two additional jets in the matrix element. The generation is performed in slices of the vector-boson  $p_T$ , and the resulting events are showered with `PYTHIA 8.2` employing a CKKW jet matching. The inclusive signal region IM3 of the analysis [4] requires  $E_T^{\text{miss}} > 350$  GeV, and for these selections the background from  $Z/W + \text{jets}$  production amounts to around 95% of the total SM background. Our  $Z/W + \text{jets}$  samples are normalised such that the different contributions match the number of events in the IM3 signal region as estimated by the ATLAS collaboration scaled to a CM energy of 14 TeV and to the appropriate integrated luminosity. In addition, the small SM backgrounds arising from  $t\bar{t}$  [23],  $tW$  [24] and diboson [25, 26] production were simulated at next-to-leading order (NLO) using `POWHEG BOX` [27]. Only the final states with at least one neutrino are considered. The samples produced with `POWHEG BOX` are normalised to the NLO cross section given by the generator, except for  $t\bar{t}$  production which is normalised to the cross section obtained at next-to-next-to-leading order plus next-to-next-to-leading logarithmic accuracy [28, 29].

The most important experimental objects in our analysis are hadronic jets and  $E_{T,\text{miss}}$ , whereas charged leptons are only used for vetoing purposes. Charged leptons are constructed from stable particles in the generator output, while jets are built by clustering the true momenta of all particles but muons that interact in the calorimeters. `FastJet` [30] is used to construct anti- $k_t$  jets [31] of radius  $R = 0.4$ . The variable  $\vec{p}_{T,\text{miss}}$  with magnitude  $E_{T,\text{miss}}$  is defined at truth level, i.e. before applying detector effects, as the vector sum of the transverse momenta of all invisible particles. The effect of the detector is simulated by applying Gaussian smearing functions to the momenta of the experimental objects and by employing reconstruction and tagging efficiency factors tuned to reproduce

	SR <sub>j</sub>	SR <sub>2j</sub>
$E_T^{\text{miss}}$	$> 350 \text{ GeV}$	
$\Delta\phi_{\vec{p}_T^{\text{miss}j}}$	$> 0.4$	
leading jet	$ \eta_{j_1}  < 2.4, \quad p_{T,j_1} > 250 \text{ GeV}$	$ \eta_{j_1}  < 2.4, \quad p_{T,j_1} > 100 \text{ GeV}$
subleading jet	$\begin{cases} \text{n/a}, & N_j = 1, \\  \eta_{j_2}  < 2.8, \quad p_{T,j_2} > 30 \text{ GeV}, & N_j > 1 \end{cases}$	$ \eta_{j_2}  < 2.8, \quad p_{T,j_2} > 50 \text{ GeV}$
$m_{j_1j_2}$	$\begin{cases} \text{n/a}, & N_j = 1, \\ < 500 \text{ GeV (800 GeV)}, & N_j > 1 \end{cases}$	$> 500 \text{ GeV (800 GeV)}$

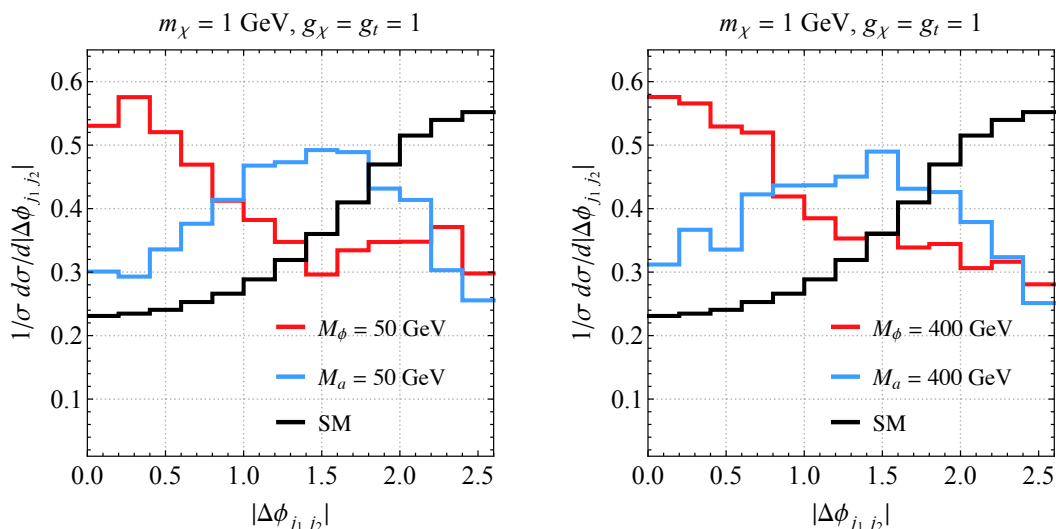
**Table 1.** Definition of the two signal regions SR<sub>j</sub> and SR<sub>2j</sub>. See the text for additional explanations.

the performance of the ATLAS detector [32, 33]. To smear  $E_T^{\text{miss}}$  the transverse momenta of unsmearred electrons, muons and jets are subtracted from the truth  $E_T^{\text{miss}}$  and replaced by the corresponding smeared quantities. The residual truth imbalance is then smeared as a function of the scalar sum of the  $p_T$  of the particles not assigned to electrons or jets. Similar techniques for fast detector simulation are used for the projection studies [34] of the ATLAS collaboration, and have also been employed in the phenomenological analyses [15, 17, 35, 36].

### 3 Analysis strategy

In our analysis, two orthogonal signal regions are defined. One focusing on the signature with a single jet (SR<sub>j</sub>), and another one with two jets of high dijet invariant mass  $m_{j_1j_2}$  (SR<sub>2j</sub>). The definitions of the signal regions are summarised in table 1. The basic selections for both signal regions require  $E_T^{\text{miss}} > 350 \text{ GeV}$ , and that the separation in the azimuthal angle  $\Delta\phi_{\vec{p}_T^{\text{miss}j}}$  between  $\vec{p}_T^{\text{miss}}$  and any jet satisfies  $\Delta\phi_{\vec{p}_T^{\text{miss}j}} > 0.4$ . Reconstructed jets have to have  $|\eta_j| < 2.8$  and  $p_{T,j} > 20 \text{ GeV}$ , and events containing more than four jets, i.e.  $N_j > 4$ , with  $p_{T,j} > 30 \text{ GeV}$  are vetoed. The latter two cuts ensure that the background from QCD multijet production is subdominant in the experimental analysis. We also veto events with electrons or muons.

Jets are treated differently in the two signal regions. In SR<sub>j</sub> (SR<sub>2j</sub>), we demand the presence of at least one jet ( $j_1$ ) with  $|\eta_{j_1}| < 2.4$  and  $p_{T,j_1} > 250 \text{ GeV}$  ( $p_{T,j_1} > 100 \text{ GeV}$ ). Notice that the  $E_T^{\text{miss}}$ , the  $\Delta\phi_{\vec{p}_T^{\text{miss}j}}$  and the leading-jet cuts imposed in SR<sub>j</sub> match those of the signal region IM3 defined in [4]. If SR<sub>j</sub> contains two or more jets we require  $|\eta_{j_2}| < 2.8$  and  $p_{T,j_2} > 30 \text{ GeV}$ , whereas for SR<sub>2j</sub> the second hardest jet has to satisfy  $|\eta_{j_2}| < 2.8$  and  $p_{T,j_2} > 50 \text{ GeV}$ . In the case of the signal region SR<sub>2j</sub>, we finally ask that the invariant mass of the two leading jets fulfills  $m_{j_1j_2} > 500 \text{ GeV}$  ( $m_{j_1j_2} > 800 \text{ GeV}$ ) in the extrapolation to  $300 \text{ fb}^{-1}$  ( $3 \text{ ab}^{-1}$ ) of integrated luminosity, while if an event in SR<sub>j</sub> features more than one jet,  $m_{j_1j_2}$  is required to be smaller than the cut employed in SR<sub>2j</sub>. We emphasise that the SR<sub>2j</sub> requirements on  $m_{j_1j_2}$  have been optimised in our analysis to provide the best possible separation between DM signal and SM background in terms of  $|\Delta\phi_{j_1j_2}|$  distributions. Such an optimisation has not been performed in the earlier study [10].



**Figure 2.** Normalised  $|\Delta\phi_{j_1j_2}|$  distributions of the  $2j + E_T^{\text{miss}}$  signature and the total SM background at the 14 TeV LHC. The shown results impose the experimental selections  $\text{SR}_{2j}$  with  $m_{j_1j_2} > 500$  GeV (cf. table 1). The choices of the spin-0  $s$ -channel DM simplified model parameters are indicated in the legends and the headlines of the two panels. For further explanations see the main text.

After applying the above cuts, the SM background amounts to approximately 595 k (102 k) events in  $\text{SR}_j$  ( $\text{SR}_{2j}$ ) per  $300\text{fb}^{-1}$  of integrated luminosity at the 14 TeV LHC. In both signal regions, the ratio of the number of DM signal to SM background events turns out to be in the ballpark of 1% (2%) for a scalar (pseudoscalar) mediator with mass below 400 GeV. For larger values of  $M_{\phi/a}$  the signal-to-background ratio rapidly decreases.

The distributions for the azimuthal angle difference between the two jets  $|\Delta\phi_{j_1j_2}|$  in the signal region  $\text{SR}_{2j}$  with  $m_{j_1j_2} > 500$  GeV are shown in figure 2. All the distributions are normalised to unity when integrated over  $|\Delta\phi_{j_1j_2}| \in [0, 2.6]$ . We see that in the case of the DM signal the normalised  $|\Delta\phi_{j_1j_2}|$  spectra display a pronounced cosine-like and sine-like modulation that, as explained for example in [10, 37, 38], is typical for scalar (red) and pseudoscalar (blue) exchange. Compared to the  $2j + E_T^{\text{miss}}$  scalar and pseudoscalar distributions the SM background (black) is peaked towards  $|\Delta\phi_{j_1j_2}| = \pi$ . The observed shape differences offer the possibility of improving the sensitivity of the mono-jet analysis and, in the case of discovery, may allow to distinguishing between a CP-even and a CP-odd mediation mechanism. It is important to notice in this respect that the shape of the  $|\Delta\phi_{j_1j_2}|$  spectra is rather insensitive to the mediator mass as illustrated by the left and right panel corresponding to masses of 50 GeV and 400 GeV, respectively. We stress that compared to the previous work [10] that used hard matrix elements only, the Monte Carlo (MC) modelling of the DM signals performed here is more sophisticated as it includes the effects of CKKW jet matching, parton shower and hadronisation corrections and a realistic detector simulation (cf. section 2). Our study thus shows that the azimuthal angle difference  $\Delta\phi_{j_1j_2}$  in  $2j + E_T^{\text{miss}}$  production is not washed out by soft physics and/or

detector effects, and therefore provides a powerful model discriminant in realistic LHC mono-jet analyses. In appendix A we quantify the gain in sensitivity that is achieved by adding shape information to the search strategy  $SR_{2j}$ .

Before discussing our LHC Run-3 and HL-LHC projections, we briefly comment on the relevance of the  $t\bar{t}$ ,  $tW$  and diboson backgrounds in our analysis. In the signal region  $SR_j$ , we find that the impact of the sum of  $t\bar{t}$  and  $tW$  (diboson) production is negligible as this contribution amounts to a fraction of only 0.2% (0.8%) of the total SM background. In  $SR_{2j}$  the contribution due to  $t\bar{t}$  and  $tW$  (diboson) production instead amounts to 6% (3%). The  $|\Delta\phi_{j_1j_2}|$  distribution of the  $t\bar{t}$  and  $tW$  background is however peaked at around  $|\Delta\phi_{j_1j_2}| = 2.8$  with a flat tail that slowly grows from  $|\Delta\phi_{j_1j_2}| = 0$ , whereas the shape of the  $|\Delta\phi_{j_1j_2}|$  spectrum of the diboson background resembles that of the leading SM background from  $Z/W + \text{jets}$  production (cf. the black histograms in figure 2). The discriminating power of the  $\Delta\phi_{j_1j_2}$  observable is therefore not affected by the subleading backgrounds if the  $\Delta\phi_{j_1j_2}$  shape fit is limited to the range  $|\Delta\phi_{j_1j_2}| \in [0, 2.6]$ , as done in the subsequent numerical analysis.

#### 4 LHC Run-3 and HL-LHC projections

The goal of this section is to derive upper limits on the signal strength  $\mu$ , i.e. the ratio of the signal yield to that predicted in the spin-0  $s$ -channel DM simplified models (2.1). Given that in both signal regions the signal-to-background ratio is at the percent level, a shape fit to a discriminant variable is necessary to maximise the sensitivity to the DM signal. In our analysis, we perform a standard  $E_T^{\text{miss}}$  shape fit in the signal region  $SR_j$ , while in  $SR_{2j}$  the shape of the  $|\Delta\phi_{j_1j_2}|$  distributions is used as a discriminator. In the case of the HL-LHC, the high number of events in the two signal regions implies that the sensitivity of the search largely depends on how well the SM background can be modelled and/or constrained. Since the systematic uncertainties plaguing the  $Z/W + \text{jets}$  background have been identified as the limiting factor in mono-jet analyses, much experimental and theoretical effort went into minimising these uncertainties by employing techniques that involve a mix of data-driven methods and MC studies (see e.g. [4–6]). Since it is beyond the scope of this work to perform such an evaluation on our MC generated SM backgrounds, we will rely on published experimental data to approximate the effect of systematic uncertainties on the LHC sensitivity perspectives.

Given that apart from the additional  $m_{j_1j_2}$  cut in the  $N_j > 1$  case, our signal region  $SR_j$  resembles the requirements of the selection IM3 of the ATLAS search [4], we base our extrapolations on the information provided in the latter article. In the recent mono-jet analysis of the ATLAS collaboration, the systematic uncertainties are evaluated through a combined shape fit to the signal region and to appropriate control regions enriched by the dominant SM backgrounds. The obtained systematic uncertainty on the number of SM background events in IM3 amounts to 2.6%, and we assume that uncertainties of the same size also arise in the case of our signal regions  $SR_j$  and  $SR_{2j}$ . Besides the systematic uncertainty on the normalisation of the SM background, also the shapes of the distributions which enter the likelihood fits carry uncertainties. In the case of [4] for instance,

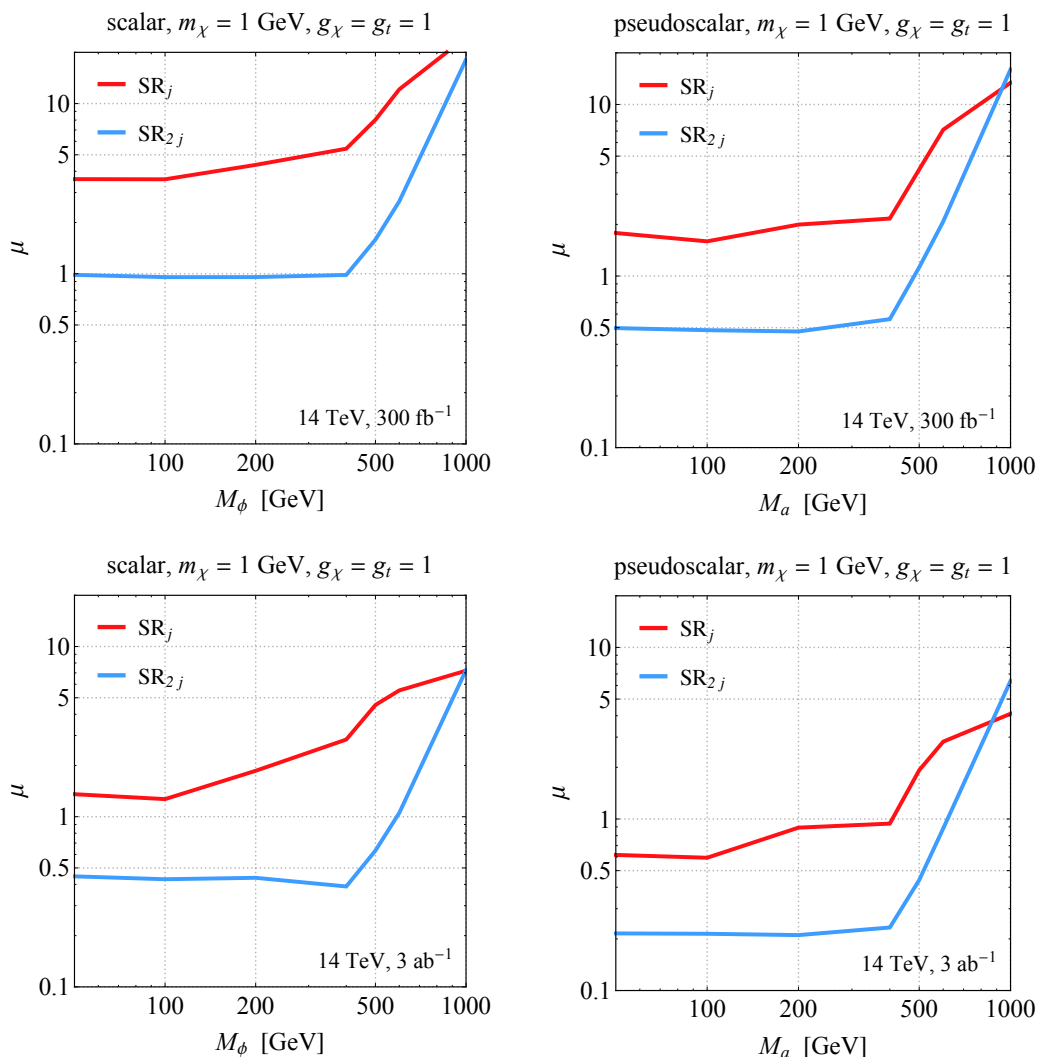
the bin-by-bin systematic uncertainties on the shape of the  $E_T^{\text{miss}}$  distribution amount to around [3%, 7%] for  $E_T^{\text{miss}} \in [350 \text{ GeV}, 1 \text{ TeV}]$ . These uncertainties are however strongly correlated among bins and cannot be naively used as bin-by-bin errors in a likelihood fit. In addition, no experimental information on the systematic uncertainties of the  $\Delta\phi_{j_1j_2}$  distributions is available, although it seems likely that the shapes of the spectra shown in figure 2 can be modelled with higher precision than the steeply falling  $E_T^{\text{miss}}$  distributions. As we are mainly interested in the relative reach of the two signal regions  $\text{SR}_j$  and  $\text{SR}_{2j}$ , as a minimal approach we only use the uncertainty of 2.6% on the total number of expected events in each signal region, ascribing no additional error to the shapes of the  $E_T^{\text{miss}}$  and  $\Delta\phi_{j_1j_2}$  distributions. This procedure will allow us to calculate an upper limit on the sensitivity of our analysis strategy.

In order to evaluate the upper confidence level (CL) limits on the signal strength  $\mu$ , we construct parametrised probability density functions for the DM signals and the SM background with the `HistFactory` package [39]. The significance is then calculated using the  $CL_s$  method [40]. The actual calculation is performed with the `Roostats` toolkit [41], which utilises the asymptotic formulas for likelihood-based tests presented in [42]. The assumptions on systematic uncertainties incorporated in the probability density functions for each of the signal regions has already been discussed before. Figure 3 displays our 95% CL limits for integrated luminosities of  $300 \text{ fb}^{-1}$  (upper row) and  $3 \text{ ab}^{-1}$  (lower row) as a function of the mediator mass for scalar (left) and pseudoscalar (right) mediators. The red (blue) curves correspond to the results of the  $E_T^{\text{miss}}$  ( $\Delta\phi_{j_1j_2}$ ) shape fit in  $\text{SR}_j$  ( $\text{SR}_{2j}$ ) as described above. The corresponding fit ranges are  $E_T^{\text{miss}} > 350 \text{ GeV}$  and  $|\Delta\phi_{j_1j_2}| \in [0, 2.6]$ , respectively. One observes that under the assumption of systematic uncertainties of 2.6% on the number of SM background events in  $\text{SR}_j$  and  $\text{SR}_{2j}$ , the  $\Delta\phi_{j_1j_2}$  shape fit proposed by us leads to significantly stronger LHC Run-3 and the HL-LHC constraints on  $\mu$  than a standard  $E_T^{\text{miss}}$  shape analysis. This finding can be understood qualitatively by recalling that there is a large fraction of two jet-like events in the case of gluon-fusion induced mono-jet production [43], and our  $\Delta\phi_{j_1j_2}$  shape fit exploits this feature. Numerically, our  $\text{SR}_{2j}$  search strategy leads to the 95% CL limits  $M_\phi > 402 \text{ GeV}$  and  $M_a > 477 \text{ GeV}$  for  $m_\chi = 1 \text{ GeV}$ ,  $g_\chi = g_t = 1$  and  $300 \text{ fb}^{-1}$  of 14 TeV data. The corresponding bounds for  $3 \text{ ab}^{-1}$  of integrated luminosity read  $M_\phi > 587 \text{ GeV}$  and  $M_a > 608 \text{ GeV}$ . We emphasise that the quoted exclusions have been obtained under the assumption that only the total number of expected events carries a systematic uncertainty. This procedure hence leads to upper bounds on the sensitivity of our analysis strategy  $\text{SR}_{2j}$ , corresponding to the limit of  $\Delta\phi_{j_1j_2}$  distributions with vanishing shape uncertainties. Since a reliable estimate of shape uncertainties and their correlations is only possible in an analysis that uses real LHC data, it is beyond the scope of this work to quantify to which extent our projections would be weakened if shape uncertainties were to be included. Additional extrapolations based on a different systematic uncertainty scenario can be found in appendix A.

## 5 Conclusions and outlook

The main goal of this article was to put the earlier study [10] of angular correlations in gluon-fusion production of loop-induced  $2j + E_T^{\text{miss}}$  signatures on more solid ground both





**Figure 3.** Value of the signal strength  $\mu$  that can be excluded at 95% CL as a function of the mass for scalar (left) and pseudoscalar (right) mediators. The reach for 300 fb<sup>-1</sup> (upper row) and 3 ab<sup>-1</sup> (lower row) of 14 TeV LHC data is shown for the parameter choices  $m_\chi = 1$  GeV and  $g_\chi = g_t = 1$ . The displayed results assume an uncertainty of 2.6% on the total number of expected events in both signal regions. Consult the text for additional explanations.

from a theoretical and experimental point of view. To this purpose, we have performed state-of-the-art MC simulations of both the mono-jet signal in spin-0  $s$ -channel simplified models and the associated SM backgrounds. The dominant background from  $Z/W +$  jets production, but also the subleading  $t\bar{t}$ ,  $tW$  and diboson channels have been considered. Our event generation includes the effects of jet matching and merging as well as parton shower and hadronisation corrections, and we have performed a realistic detector modelling (see section 2).

The proposed analysis strategy aims at separating the mono-jet signature into two distinct signal regions. The first signal region called SR<sub>j</sub> focuses on single jet-like events, while the second signal region referred to as SR<sub>2j</sub> requires the presence of a jet pair with

large invariant mass in the final state (see section 3). In the signal region  $\text{SR}_{2j}$ , we have studied the azimuthal angle difference  $\Delta\phi_{j_1j_2}$  in mono-jet production. Our study shows that the latter observable provides a powerful model discriminant in realistic LHC analyses. In fact, shape fits to the  $\Delta\phi_{j_1j_2}$  variable will generically help to improve the sensitivity of mono-jet searches (see appendix A), and, in the case of a discovery, might allow to distinguishing between scalar and pseudoscalar mediators.

We have then analysed the mono-jet coverage of the parameter space of the spin-0  $s$ -channel DM simplified models expected at LHC Run-3 and the HL-LHC (see section 4). In a first step, we have derived hypothetical limits on the signal strength that follow from a standard  $E_T^{\text{miss}}$  shape analysis to the  $\text{SR}_j$  signal region. In a second step, we have then obtained bounds by performing shape fits to the  $\Delta\phi_{j_1j_2}$  variable utilising the  $\text{SR}_{2j}$  event samples. Under the reasonable assumption that the systematic uncertainties on the number of SM background events in  $\text{SR}_j$  and  $\text{SR}_{2j}$  are the same, we have found that the proposed  $\Delta\phi_{j_1j_2}$  shape fit has a significantly better reach than a standard  $E_T^{\text{miss}}$  shape analysis. For the benchmark parameter choices  $m_\chi = 1 \text{ GeV}$  and  $g_\chi = g_t = 1$ , the 95% CL exclusion limits that derive from our  $\text{SR}_{2j}$  search strategy read  $M_\phi > 402 \text{ GeV}$  and  $M_a > 477 \text{ GeV}$  for  $300 \text{ fb}^{-1}$  of 14 TeV data. The corresponding bounds for  $3 \text{ ab}^{-1}$  of integrated luminosity turn out to be  $M_\phi > 587 \text{ GeV}$  and  $M_a > 608 \text{ GeV}$ . Notice that in the scalar case the quoted LHC Run-3 limit is slightly weaker than the bound that follows from a combined analysis of  $t\bar{t} + E_T^{\text{miss}}$  and  $tW + E_T^{\text{miss}}$  production [17], while in all other cases the mono-jet sensitivity exceeds that of the  $tX + E_T^{\text{miss}}$  search. This finding illustrated the synergy and complementarity of the latter two mono- $X$  channels [44, 45] in the context of spin-0  $s$ -channel DM simplified models.

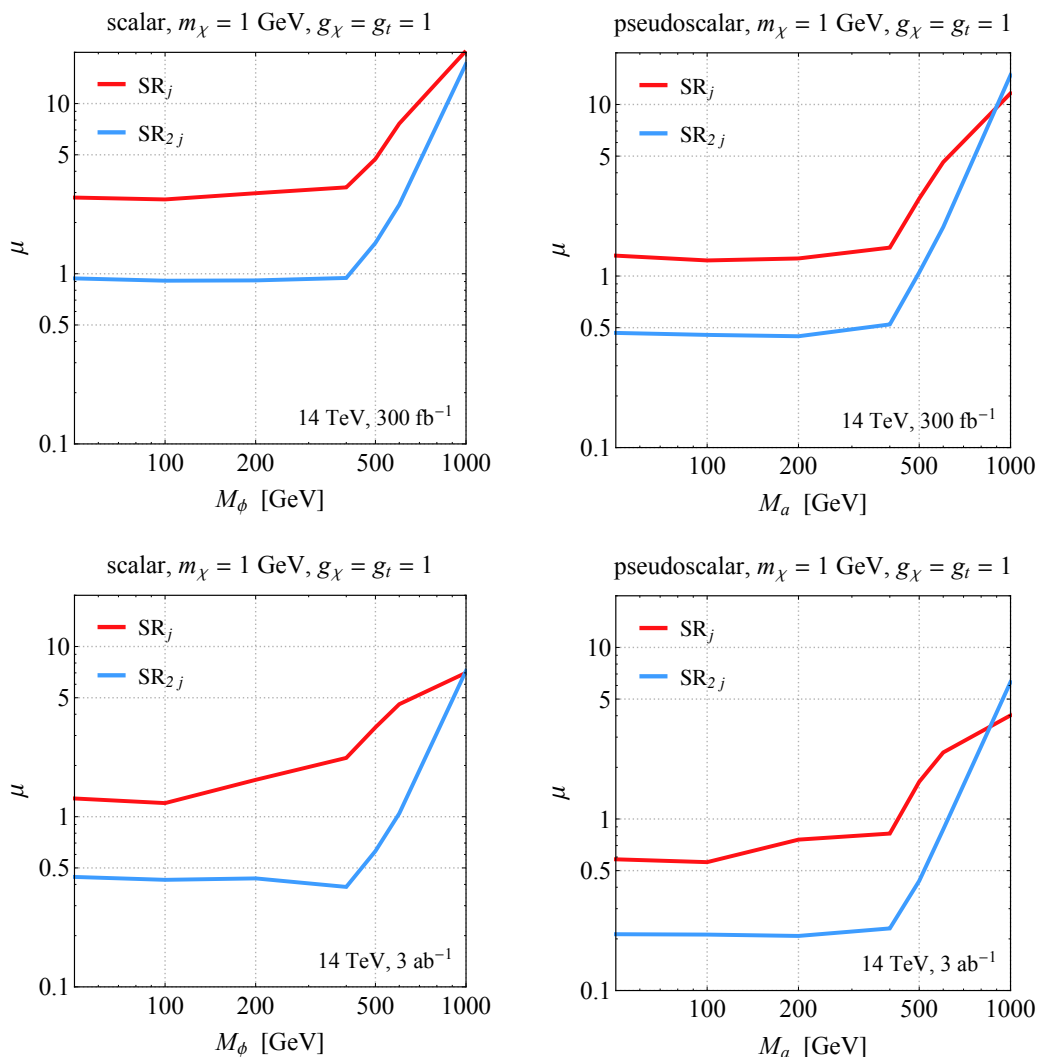
We finally note that the  $\text{SR}_{2j}$  analysis strategy proposed by us can also be straightforwardly applied to next-generation DM simplified models such two-Higgs-doublet extensions with an extra spin-0 gauge singlet [46–52]. Like in the case of the spin-0  $s$ -channel DM simplified models discussed here, we expect that exploiting the  $\Delta\phi_{j_1j_2}$  correlations in  $2j + E_T^{\text{miss}}$  production will also allow to significantly strengthen future LHC mono-jet constraints on spin-0 next-generation DM simplified models.

## Acknowledgments

We are grateful to Giuliano Gustavino for useful comments on the manuscript.

## A Supplementary material

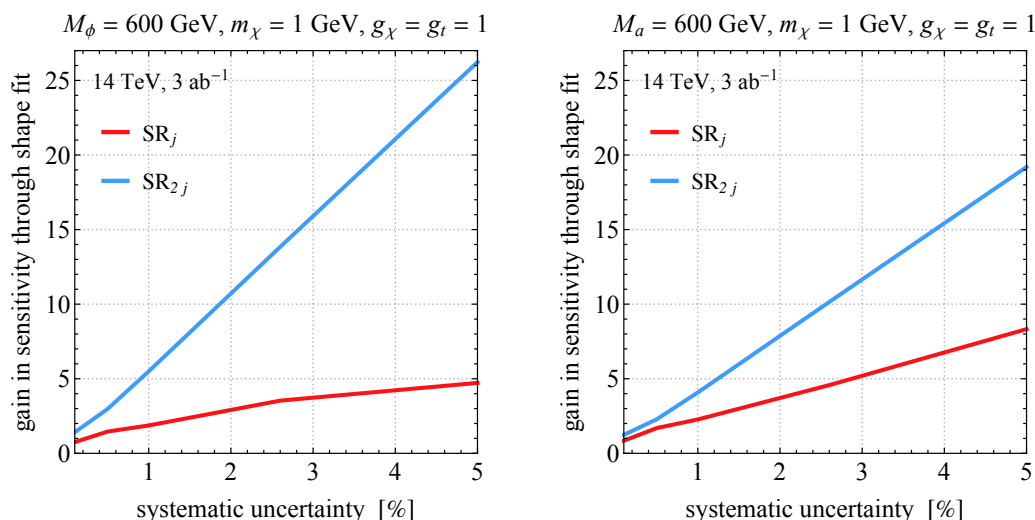
In this appendix we extend the numerical study performed in section 4. We start by presenting LHC explorations based on an alternative more aggressive assumption about the systematic uncertainties of future LHC mono-jet searches. Anticipating improvements in detector performance and modelling of DM signal and SM background processes, we assume, in the spirit of [53, 54], that the present systematic uncertainties on the total number of expected events in the signal regions  $\text{SR}_j$  and  $\text{SR}_{2j}$  can be reduced by a factor of 2. In figure 4 we show the 95% CL limits for  $300 \text{ fb}^{-1}$  (upper row) and  $3 \text{ ab}^{-1}$  (lower row)



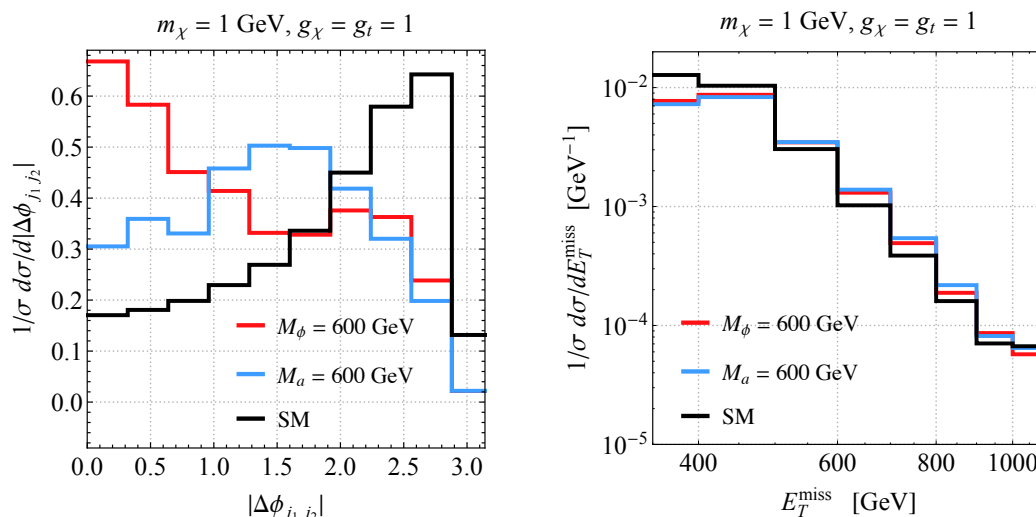
**Figure 4.** As figure 3 but assuming uncertainties of 1.3% on the total number of expected events in the signal regions  $SR_j$  and  $SR_{2j}$ . See the text for further explanations.

of data as a function of the scalar (left) and pseudoscalar (right) mediator mass. The red (blue) curves illustrate the results of the  $E_T^{\text{miss}} (\Delta\phi_{j_1 j_2})$  shape fit in  $SR_j$  ( $SR_{2j}$ ) as described in section 4, assuming an improved systematic uncertainty of 1.3%. Under this assumption, we find that the proposed  $SR_{2j}$  search strategy leads to the 95% CL limits  $M_\phi > 409$  GeV and  $M_a > 490$  GeV for  $m_\chi = 1$  GeV,  $g_\chi = g_t = 1$  and  $300 \text{ fb}^{-1}$  of 14 TeV data. The corresponding  $3 \text{ ab}^{-1}$  bounds are  $M_\phi > 589$  GeV and  $M_a > 609$  GeV. Notice that these limits are only marginally better than the bounds reported at the end of section 4. The 95% CL bounds on  $\mu$  that derive from the search strategy  $SR_j$  are in contrast notable improved if the systematic uncertainties are reduced from 2.6% to 1.3%. Numerically, we find average improvements of 45% and 15% at LHC Run-3 and HL-LHC, respectively.

In addition let us quantify the impact of shape information in the two mono-jet search strategies considered by us. To do so, we define the gain of sensitivity through the shape fit



**Figure 5.** Gain in sensitivity that is achieved by adding shape information to the search strategy  $SR_j$  (red curves) and  $SR_{2j}$  (blue curves) as a function of the assumed systematic uncertainties in percent. The shown results correspond to  $3 \text{ ab}^{-1}$  of 14 TeV LHC data and the used parameter choices are indicated in the headlines of the two panels. See text for additional information.



**Figure 6.** Normalised  $|\Delta\phi_{j_1 j_2}|$  (left) and  $E_T^{\text{miss}}$  (right) distributions in the  $SR_{2j}$  and  $SR_j$ , respectively. The last bin of the  $E_T^{\text{miss}}$  histograms is an overflow bin. The shown results correspond to  $3 \text{ ab}^{-1}$  of 14 TeV data and they impose the HL-LHC cuts as specified in table 1. The red (blue) histograms represent the DM signal arising from scalar (pseudoscalar) exchange, while the SM background distributions are coloured black. The legends and the headlines of the panels indicate the used spin-0  $s$ -channel DM simplified model parameters.

as the ratio of  $\mu$  values obtained with and without the inclusion of shape information. This ratio is displayed in figure 5 as a function of the assumed systematic uncertainty on the number of events in  $SR_j$  (red curves) and  $SR_{2j}$  (blue curves). The shown results correspond to the HL-LHC and two benchmark spin-0  $s$ -channel DM simplified models. From the

panels it is evident that the shape information carried by  $\Delta\phi_{j_1j_2}$  is a significantly more powerful constraint than that of  $E_T^{\text{miss}}$ . This finding is unsurprising, if one considers the shapes of the  $\Delta\phi_{j_1j_2}$  and  $E_T^{\text{miss}}$  corresponding to the parameter choices used to obtain the latter figure. As can be seen from figure 6, the  $\Delta\phi_{j_1j_2}$  spectrum displays a marked cosine-like (sine-like) modulation in the scalar (pseudoscalar) case, while the  $E_T^{\text{miss}}$  distributions are steeply falling and largely independent of the mediator type. In the case of the  $\Delta\phi_{j_1j_2}$  distributions, one furthermore observes a clear distinction between the shapes of the DM signals and the SM background, while in the  $E_T^{\text{miss}}$  case the differences between the three normalised spectra are significantly less prominent.

The features of the results shown in figures 3, 4, 5 and 6 thus strongly suggest that search strategies based on  $\Delta\phi_{j_1j_2}$  shape fits are not only more powerful than standard  $E_T^{\text{miss}}$  shape analyses in constraining the parameter space of spin-0  $s$ -channel DM simplified models, but are also less dependent on hypothetical improvements of the systematic uncertainties of future mono-jet searches.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] UA1 collaboration, *Experimental observation of events with large missing transverse energy accompanied by a jet or a photon(s) in  $p\bar{p}$  collisions at  $\sqrt{s} = 540$  GeV*, *Phys. Lett. B* **139** (1984) 115.
- [2] D0 collaboration, *Search for large extra dimensions in the monojet + missing  $E_T$  channel at  $D\bar{O}$* , *Phys. Rev. Lett.* **90** (2003) 251802 [[hep-ex/0302014](https://arxiv.org/abs/hep-ex/0302014)] [[INSPIRE](#)].
- [3] CDF collaboration, *Search for large extra dimensions in final states containing one photon or jet and large missing transverse energy produced in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV*, *Phys. Rev. Lett.* **101** (2008) 181602 [[arXiv:0807.3132](https://arxiv.org/abs/0807.3132)] [[INSPIRE](#)].
- [4] ATLAS collaboration, *Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector*, *JHEP* **01** (2018) 126 [[arXiv:1711.03301](https://arxiv.org/abs/1711.03301)] [[INSPIRE](#)].
- [5] CMS collaboration, *Search for new physics in final states with an energetic jet or a hadronically decaying  $W$  or  $Z$  boson and transverse momentum imbalance at  $\sqrt{s} = 13$  TeV*, *Phys. Rev. D* **97** (2018) 092005 [[arXiv:1712.02345](https://arxiv.org/abs/1712.02345)] [[INSPIRE](#)].
- [6] J.M. Lindert et al., *Precise predictions for  $V + \text{jets}$  dark matter backgrounds*, *Eur. Phys. J. C* **77** (2017) 829 [[arXiv:1705.04664](https://arxiv.org/abs/1705.04664)] [[INSPIRE](#)].
- [7] J. Abdallah et al., *Simplified models for dark matter searches at the LHC*, *Phys. Dark Univ.* **9-10** (2015) 8 [[arXiv:1506.03116](https://arxiv.org/abs/1506.03116)] [[INSPIRE](#)].
- [8] D. Abercrombie et al., *Dark matter benchmark models for early LHC run-2 searches: report of the ATLAS/CMS dark matter forum*, [arXiv:1507.00966](https://arxiv.org/abs/1507.00966) [[INSPIRE](#)].
- [9] G. Busoni et al., *Recommendations on presenting LHC searches for missing transverse energy signals using simplified  $s$ -channel models of dark matter*, [arXiv:1603.04156](https://arxiv.org/abs/1603.04156) [[INSPIRE](#)].

- [10] U. Haisch, A. Hibbs and E. Re, *Determining the structure of dark-matter couplings at the LHC*, *Phys. Rev. D* **89** (2014) 034009 [[arXiv:1311.7131](#)] [[INSPIRE](#)].
- [11] O.J.P. Eboli and D. Zeppenfeld, *Observing an invisible Higgs boson*, *Phys. Lett. B* **495** (2000) 147 [[hep-ph/0009158](#)] [[INSPIRE](#)].
- [12] R.C. Cotta, J.L. Hewett, M.P. Le and T.G. Rizzo, *Bounds on dark matter interactions with electroweak gauge bosons*, *Phys. Rev. D* **88** (2013) 116009 [[arXiv:1210.0525](#)] [[INSPIRE](#)].
- [13] A. Crivellin, U. Haisch and A. Hibbs, *LHC constraints on gauge boson couplings to dark matter*, *Phys. Rev. D* **91** (2015) 074028 [[arXiv:1501.00907](#)] [[INSPIRE](#)].
- [14] M.R. Buckley and D. Goncalves, *Constraining the strength and CP structure of dark production at the LHC: the associated top-pair channel*, *Phys. Rev. D* **93** (2016) 034003 [[arXiv:1511.06451](#)] [[INSPIRE](#)].
- [15] U. Haisch, P. Pani and G. Polesello, *Determining the CP nature of spin-0 mediators in associated production of dark matter and  $t\bar{t}$  pairs*, *JHEP* **02** (2017) 131 [[arXiv:1611.09841](#)] [[INSPIRE](#)].
- [16] G. Bélanger, R.M. Godbole, C.K. Khosa and S.D. Rindani, *Probing CP nature of a mediator in associated production of dark matter with single top quark*, [arXiv:1811.11048](#) [[INSPIRE](#)].
- [17] U. Haisch and G. Polesello, *Searching for production of dark matter in association with top quarks at the LHC*, *JHEP* **02** (2019) 029 [[arXiv:1812.00694](#)] [[INSPIRE](#)].
- [18] M. Backović et al., *Higher-order QCD predictions for dark matter production at the LHC in simplified models with s-channel mediators*, *Eur. Phys. J. C* **75** (2015) 482 [[arXiv:1508.05327](#)] [[INSPIRE](#)].
- [19] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [[arXiv:1405.0301](#)] [[INSPIRE](#)].
- [20] R.D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [[arXiv:1207.1303](#)] [[INSPIRE](#)].
- [21] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[arXiv:1410.3012](#)] [[INSPIRE](#)].
- [22] S. Catani, F. Krauss, R. Kuhn and B.R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [[hep-ph/0109231](#)] [[INSPIRE](#)].
- [23] J.M. Campbell, R.K. Ellis, P. Nason and E. Re, *Top-pair production and decay at NLO matched with parton showers*, *JHEP* **04** (2015) 114 [[arXiv:1412.1828](#)] [[INSPIRE](#)].
- [24] E. Re, *Single-top  $Wt$ -channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547 [[arXiv:1009.2450](#)] [[INSPIRE](#)].
- [25] T. Melia, P. Nason, R. Rontsch and G. Zanderighi,  *$W^+W^-$ ,  $WZ$  and  $ZZ$  production in the POWHEG BOX*, *JHEP* **11** (2011) 078 [[arXiv:1107.5051](#)] [[INSPIRE](#)].
- [26] P. Nason and G. Zanderighi,  *$W^+W^-$ ,  $WZ$  and  $ZZ$  production in the POWHEG-BOX-V2*, *Eur. Phys. J. C* **74** (2014) 2702 [[arXiv:1311.1365](#)] [[INSPIRE](#)].
- [27] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [[arXiv:1002.2581](#)] [[INSPIRE](#)].

- [28] M. Czakon and A. Mitov, *Top++: a program for the calculation of the top-pair cross-section at hadron colliders*, *Comput. Phys. Commun.* **185** (2014) 2930 [[arXiv:1112.5675](#)] [[INSPIRE](#)].
- [29] M. Czakon, P. Fiedler and A. Mitov, *Total top-quark pair-production cross section at hadron colliders through  $O(\alpha_S^4)$* , *Phys. Rev. Lett.* **110** (2013) 252004 [[arXiv:1303.6254](#)] [[INSPIRE](#)].
- [30] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [[arXiv:1111.6097](#)] [[INSPIRE](#)].
- [31] M. Cacciari, G.P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)] [[INSPIRE](#)].
- [32] ATLAS collaboration, *The ATLAS experiment at the CERN Large Hadron Collider*, *2008 JINST* **3** S08003 [[INSPIRE](#)].
- [33] ATLAS collaboration, *Expected performance of the ATLAS experiment — Detector, trigger and physics*, [arXiv:0901.0512](#) [[INSPIRE](#)].
- [34] ATLAS collaboration, *Expected performance for an upgraded ATLAS detector at High-Luminosity LHC*, *ATL-PHYS-PUB-2016-026* (2016).
- [35] P. Pani and G. Polesello, *Dark matter production in association with a single top-quark at the LHC in a two-Higgs-doublet model with a pseudoscalar mediator*, *Phys. Dark Univ.* **21** (2018) 8 [[arXiv:1712.03874](#)] [[INSPIRE](#)].
- [36] U. Haisch and G. Polesello, *Searching for heavy Higgs bosons in the  $t\bar{t}Z$  and  $tbW$  final states*, *JHEP* **09** (2018) 151 [[arXiv:1807.07734](#)] [[INSPIRE](#)].
- [37] T. Plehn, D.L. Rainwater and D. Zeppenfeld, *Determining the structure of Higgs couplings at the LHC*, *Phys. Rev. Lett.* **88** (2002) 051801 [[hep-ph/0105325](#)] [[INSPIRE](#)].
- [38] G. Klamke and D. Zeppenfeld, *Higgs plus two jet production via gluon fusion as a signal at the CERN LHC*, *JHEP* **04** (2007) 052 [[hep-ph/0703202](#)] [[INSPIRE](#)].
- [39] ROOT collaboration *HistFactory: a tool for creating statistical models for use with RooFit and RooStats*, *CERN-OPEN-2012-016* (2012).
- [40] A.L. Read, *Presentation of search results: the  $CL_s$  technique*, *J. Phys. G* **28** (2002) 2693 [[INSPIRE](#)].
- [41] L. Moneta et al., *The RooStats Project*, *PoS(ACAT2010)057* [[arXiv:1009.1003](#)] [[INSPIRE](#)].
- [42] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [*Erratum ibid.* **C 73** (2013) 2501] [[arXiv:1007.1727](#)] [[INSPIRE](#)].
- [43] U. Haisch, F. Kahlhoefer and E. Re, *QCD effects in mono-jet searches for dark matter*, *JHEP* **12** (2013) 007 [[arXiv:1310.4491](#)] [[INSPIRE](#)].
- [44] M.R. Buckley, D. Feld and D. Goncalves, *Scalar simplified models for dark matter*, *Phys. Rev. D* **91** (2015) 015017 [[arXiv:1410.6497](#)] [[INSPIRE](#)].
- [45] U. Haisch and E. Re, *Simplified dark matter top-quark interactions at the LHC*, *JHEP* **06** (2015) 078 [[arXiv:1503.00691](#)] [[INSPIRE](#)].
- [46] S. Ipek, D. McKeen and A.E. Nelson, *A renormalizable model for the galactic center gamma ray excess from dark matter annihilation*, *Phys. Rev. D* **90** (2014) 055021 [[arXiv:1404.3716](#)] [[INSPIRE](#)].

- [47] J.M. No, *Looking through the pseudoscalar portal into dark matter: novel mono-Higgs and mono-Z signatures at the LHC*, *Phys. Rev. D* **93** (2016) 031701 [[arXiv:1509.01110](#)] [[INSPIRE](#)].
- [48] D. Goncalves, P.A.N. Machado and J.M. No, *Simplified models for dark matter face their consistent completions*, *Phys. Rev. D* **95** (2017) 055027 [[arXiv:1611.04593](#)] [[INSPIRE](#)].
- [49] N.F. Bell, G. Busoni and I.W. Sanderson, *Self-consistent Dark Matter Simplified Models with an s-channel scalar mediator*, *JCAP* **03** (2017) 015 [[arXiv:1612.03475](#)] [[INSPIRE](#)].
- [50] M. Bauer, U. Haisch and F. Kahlhoefer, *Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators*, *JHEP* **05** (2017) 138 [[arXiv:1701.07427](#)] [[INSPIRE](#)].
- [51] P. Tunney, J.M. No and M. Fairbairn, *Probing the pseudoscalar portal to dark matter via  $\bar{b}bZ(\rightarrow \ell\ell)+\cancel{E}_T$ : from the LHC to the Galactic Center excess*, *Phys. Rev. D* **96** (2017) 095020 [[arXiv:1705.09670](#)] [[INSPIRE](#)].
- [52] LHC DARK MATTER WORKING GROUP collaboration, *LHC dark matter working group: next-generation spin-0 dark matter models*, [arXiv:1810.09420](#) [[INSPIRE](#)].
- [53] CMS Collaboration, *Estimated sensitivity for new particle searches at the HL-LHC*, [CMS-PAS-FTR-16-005](#) (2016).
- [54] ATLAS collaboration, *Extrapolation of  $E_T^{miss} + jet$  search results to an integrated luminosity of  $300\text{ fb}^{-1}$  and  $3000\text{ fb}^{-1}$* , [ATL-PHYS-PUB-2018-043](#) (2018).