Search for long-lived neutral particles in $pp$ collisions at $\sqrt{s} = 13$ TeV that decay into displaced hadronic jets in the ATLAS calorimeter

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 11 February 2019 / Accepted: 18 May 2019
© CERN for the benefit of the ATLAS collaboration 2019

Abstract This paper describes a search for pairs of neutral, long-lived particles decaying in the ATLAS calorimeter. Long-lived particles occur in many extensions to the Standard Model and may elude searches for new promptly decaying particles. The analysis considers neutral, long-lived scalars with masses between 5 and 400 GeV, produced from decays of heavy bosons with masses between 125 and 1000 GeV, where the long-lived scalars decay into Standard Model fermions. The analysis uses either 10.8 fb$^{-1}$ or 33.0 fb$^{-1}$ of data (depending on the trigger) recorded in 2016 at the LHC with the ATLAS detector in proton–proton collisions at a centre-of-mass energy of 13 TeV. No significant excess is observed, and limits are reported on the production cross section times branching ratio as a function of the proper decay length of the long-lived particles.

1 Introduction

Long-lived particles (LLPs) feature in a variety of models that have been proposed to address some of the open questions of the Standard Model (SM). Examples are: various supersymmetric (SUSY) models [1–7]; Neutral Naturalness [8–11] and Hidden Sector (HS) [12–14] models that address the hierarchy problem; models that seek to incorporate dark matter [15–18], or explain the matter–antimatter asymmetry of the universe [19]; and models that lead to massive neutrinos [20,21]. Decays of LLPs created in collider experiments would produce unique signatures that may have been overlooked by previous searches for particles that decay promptly. This paper presents a search sensitive to neutral LLPs decaying mainly in the hadronic calorimeter (HCal) or at the outer edge of the electromagnetic calorimeter (ECal) of the ATLAS detector. This allows the analysis to probe LLP proper decay lengths ($c\tau$, where $c$ is the speed of light and $\tau$ is the lifetime of the LLP) ranging between a few centimetres and a few tens of metres. In HS models, a proposed new set of particles and forces is weakly coupled to the SM via a mediator particle. As a benchmark, this analysis uses a simplified HS model [12–14,22,23], in which the SM and HS are connected via a heavy neutral boson ($\Phi$), which may decay into two long-lived neutral scalar bosons ($s$). The neutral scalars are assumed not to interact with the detector. While $\Phi$ could be the Higgs boson, this analysis considers mediators with masses ranging from 125 to 1000 GeV, and scalars with masses between 5 and 400 GeV. The decay $\Phi \rightarrow ss \rightarrow f \bar{f} f' \bar{f}'$ is considered, where $f$ refers to fermions. Decays to bosons are not considered in the benchmark model used in this analysis. Since this model assumes that the branching ratios of the scalar decaying into SM fermions are the same as those of the SM Higgs, each long-lived scalar usually decays into heavy fermions: $b\bar{b}$, $c\bar{c}$, and $\tau^+\tau^-$. The branching ratio among the different decays depends on the mass of the scalar but for $m_s \geq 25$ GeV it is almost constant and equal to 85:5:8. The SM quarks from the LLP decay hadronize, resulting in jets whose origins may be far from the interaction point (IP) of the collision. The proper decay lengths of LLPs in HS models are typically unconstrained, aside from a rough upper limit of $c\tau \lesssim 10^8$ m given by the cosmological constraint of Big Bang Nucleosynthesis [24], and could be short enough for the LLPs to decay inside the ATLAS detector volume.

Previous searches for pair-produced neutral LLPs at hadron colliders have been performed at the Tevatron and at the LHC. At the Tevatron, searches by DØ [25] and CDF [26] looked for displaced vertices in their tracking system only, allowing them to set limits on LLP proper decay lengths of the order of a few centimetres. At the LHC, the CMS experiment has performed searches at centre-of-mass energies of 7, 8 or 13 TeV for neutral LLPs by considering events with either converted photons and missing energy [27,28], or with lepton [29,30] or jet pairs [31,32] originating from displaced vertices in the tracking system. A CMS search for jet pairs originating in the tracker was also performed at 13 TeV [33]. The CMS searches are sensitive to LLP proper decay lengths
from ~0.1 mm to ~2 m. Previous ATLAS searches for neutral LLPs consider events with photons [34], or particles originating from displaced vertices in the tracking system [35,36]. Other searches involve pairs of displaced jets in the HCal (8 TeV) [37,38], or pairs of reconstructed vertices in the muon spectrometer (MS) at 7 and 13 TeV [39,40], or the combination of one displaced vertex in the MS and one in the inner tracking detector (8 TeV) [41]. Other searches consider pairs of muons originating after the inner tracker [42,43]. These ATLAS searches are complementary, since they use different subdetectors, and therefore their sensitivities are governed by different instrumental effects and sub-detector responses to the kinematics of the LLP decays. They also have different backgrounds, and different lifetime coverage due to the different physical location of the sub-detectors, with sensitivity to LLP proper decay lengths extending from a few millimeters to about 200 m.

The analysis presented in this paper is an update to the 8 TeV ATLAS search for pair-produced neutral LLPs decaying in the HCal [37], using 10.8 fb$^{-1}$ or 33.0 fb$^{-1}$ of 13 TeV data depending on the trigger, with significant improvements to the displaced-jet identification, event selection and background estimation. If the scalar decay occurs in the calorimeters, the two resulting quarks are reconstructed as a single jet with unusual features compared to jets from SM processes. These jets will typically have no associated activity in the tracking system. Furthermore, they will often have a high ratio of energy deposited in the HCal ($E_H$) to energy deposited in the ECAL ($E_{EM}$). This ratio, $E_H/E_{EM}$, is referred to as the CalRatio. Finally, jets resulting from these decays will appear narrower than prompt jets when reconstructed with standard algorithms. This analysis requires two such non-standard jets.

The main background process that mimics this signature is SM multijet production, in cases where the jets are composed mainly of neutral hadrons or are mis-reconstructed due to noise or instrumental effects. Despite the low probability of a prompt jet to produce a signal-like jet, the SM multijet rate is high enough for this to be the dominant background. Other contributions come from the non-collision background consisting of cosmic rays and beam-induced background (BIB) [44]. The latter is composed of LHC beam–gas interactions and beam-halo interactions with the collimators upstream of the ATLAS detector, resulting in muons travelling parallel to the beam-pipe.

Two triggers were used to collect the data, one optimal for models with $m_{\Phi} > 200$ GeV and the other for $m_{\Phi} \leq 200$ GeV, and different selections are used to analyse the dataset collected with each trigger. Jets are classified as signal- or background-like jets using machine learning in two steps: first, for every reconstructed jet, a multilayer perceptron, trained on signal jets from LLP decays, is used to predict the decay position of the particle that generated it; next, a per-jet Boosted Decision Tree (BDT) classifies jets as signal-like, multijet-like or BIB-like jets. Events are then classified as likely to have been produced by a signal process or a background process using a per-event BDT. Two separate versions of the per-event BDT are trained: one optimised for models with $m_{\Phi} \leq 200$ GeV (referred to as low-$m_{\Phi}$ models), and the other for models with $m_{\Phi} > 200$ GeV (high-$m_{\Phi}$ models). The final sample is constructed by making a selection on the relevant per-event BDT output value of candidate events and imposing event quality criteria and requirements to suppress cosmic rays and BIB. These selections remove almost all the non-collision background, leaving only multi-jet background, and maximise signal-to-background ratio in the final search region.

The ATLAS detector is described in Sect. 2. The collection of the data and generation of samples of simulated events are then discussed in Sect. 3. The trigger and event selection are detailed in Sect. 4, followed by a discussion of the estimate of the background yield in the search regions in Sect. 5. The systematic uncertainties are summarised in Sect. 6. The statistical interpretation of the data and combination of results with the MS displaced vertex search are described in Sect. 7, and the conclusions are given in Sect. 8.

2 ATLAS detector

The ATLAS detector [45] at the LHC covers nearly the entire solid angle around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets. The inner-detector system is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$.

The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track. The layer closest to the interaction point is known as the insertable B-layer [46–48]. It was added in 2014 and provides high-resolution hits at small radius to improve the tracking performance. The pixel detector is surrounded by the silicon microstrip tracker, which usually provides four three-dimensional measurement points per track. These silicon detectors are complemented by the transition radiation tracker, with coverage up to $|\eta| = 2.0$, which enables radially extended track reconstruction in this region.

$1$ ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upwards. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{\Delta \phi^2 + (\Delta \eta)^2}$. 

$\copyright$ Springer
The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) electromagnetic calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$, to correct for energy loss in material upstream of the calorimeters. The ECal extends from 1.5 to 2.0 m in radial distance $r$ in the barrel and from 3.6 to 4.25 m in $|z|$ in the endcaps. Hadronic calorimetry is provided by a steel/scintillator tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters covering $|\eta| > 1.5$. The HCal covers the region from 2.25 to 4.25 m in $r$ in the barrel (although the HCal active material extends only up to 3.9 m) and from 4.3 to 6.05 m in $|z|$ in the endcaps. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements respectively.

The calorimeters have a highly granular lateral and longitudinal segmentation. Including the presamplers, there are seven sampling layers in the combined central calorimeters (the LAr presampler, three in the ECal barrel and three in the HCal barrel) and eight sampling layers in the endcap region (the presampler, three in ECal endcaps and four in HCal endcaps). The forward calorimeter modules provide three sampling layers in the forward region. The total amount of material in the ECal corresponds to 24–35 radiation lengths in the barrel and 35–40 radiation lengths in the endcaps. The combined depth of the calorimeters for hadronic energy measurements is more than 9 hadronic interaction lengths nearly everywhere across the full detector acceptance.

The muon spectrometer comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in the magnetic field generated by the superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 T m (Tesla x metre) across most of the detector.

The ATLAS detector selects events using a tiered trigger system [49]. The level-1 trigger is implemented in custom electronics and reduces the event rate from the LHC crossing frequency of 40 MHz to a design value of 100 kHz. The second level, known as the high-level trigger, is implemented in software running on a commodity PC farm that processes the events and reduces the rate of recorded events to 1 kHz.

### 3 Data and simulation samples

#### 3.1 Data samples

The data used in this analysis were collected by the ATLAS detector during 2016 data-taking using proton–proton ($pp$) collisions at $\sqrt{s} = 13$ TeV. Four datasets are defined according to the trigger used to select them. The search is performed on the so-called main dataset, collected by two different LLP signature-driven triggers, referred to as the low-$E_T$ CalRatio trigger and high-$E_T$ CalRatio trigger, which are described in detail in Sect. 4. The high-$E_T$ CalRatio trigger was active during the full 2016 data-taking period. After requirements based on beam and detector conditions and data quality are applied, the data collected with this trigger corresponds to an integrated luminosity of 33.0 fb$^{-1}$. The low-$E_T$ CalRatio trigger was activated in September 2016, collecting data corresponding to an integrated luminosity of 10.8 fb$^{-1}$. The events collected with these triggers are referred to as high-$E_T$ and low-$E_T$ datasets respectively. Two additional datasets, referred to as the BIB and cosmics datasets, were collected using dedicated triggers running in special conditions, as described in Sect. 4.

#### 3.2 Signal and background simulation

The $\Phi \rightarrow ss$ signal samples were generated using MADGRAPH5 [50] at leading order (LO) with the NNPDF2.3LO parton distribution function (PDF) set [51]. The shower process was implemented using PYTHIA 8.210 [52] using the A14 set of tuned parameters (tune) [53]. Several sets of samples were generated, each modelling different combinations of $m_\Phi$ and $m_s$, with $m_\Phi \in [125, 1000]$ GeV and $m_s \in [5, 400]$ GeV. For consistency with the rest of the samples, in the $m_s = 400$ GeV case, top-quark decays were not included in the generation process, even though they are kinematically allowed. The simplified model used in the generation does not give a specific prediction for the absolute production cross section. Each sample was generated for two assumptions about the LLP decay length: one sample is used to study the signal throughout the analysis, while the other sample (with the alternate decay length assumption) is used in the training of the BDTs as well as to validate the procedure for extrapolating limits to different proper decay lengths of the long-lived scalar $s$.

The main SM background in this analysis is multijet production. Although a data-driven method is used to perform the background estimation, simulated multijet events are needed for BDT training and evaluation of some of the systematic uncertainties. The samples were generated with PYTHIA 8.186 [54] using the A14 tune for parton showering and hadronisation. The NNPDF2.3LO PDF set was used.

To model the effect of multiple $pp$ interactions in the same or neighbouring bunches (pile-up), simulated inclusive $pp$ events were overlaid on each generated signal and background event. The multiple interactions were simulated with PYTHIA 8.186 using the A2 tune [55] and the MSTW2008LO PDF set [56].

The detector response to the simulated events was evaluated with the GEANT4-based detector simulation [57, 58].
full simulation of all the detector components was used for all the samples. The standard ATLAS reconstruction software was used for both simulation and pp data.

4 Trigger and event selection

Events are first selected by two dedicated signature-driven triggers called CalRatio triggers [59], which are designed to identify jets that result from neutral LLPs decaying near the outer radius of the ECal or within the HCal. The triggers make use of the three main characteristics of the displaced jets: they are narrow jets with a high fraction of their energy deposited in the HCal and typically have no tracks pointing towards the jet. Two trigger paths are followed in this analysis, defined by two CalRatio triggers that differ only in the level-1 (L1) trigger selection. The high-$E_T$ trigger was originally designed for LHC Run 1. The trigger definition was adapted to the Run 2 higher energy and pile-up conditions by, among other modifications, raising the transverse energy ($E_T$) threshold as specified below. This higher threshold has a negative impact on the efficiency for models with $m_\phi \leq 200$ GeV. To recover efficiency for those models, a new trigger, called the low-$E_T$ trigger, was designed with a lower threshold.

At L1, the high-$E_T$ trigger selects narrow jets which each deposit $E_T > 60$ GeV in a $0.2 \times 0.2$ ($\Delta \eta \times \Delta \phi$) region of the ECal and HCal combined [60]. In September 2016 an upgraded L1 trigger component, the topological trigger, was commissioned in ATLAS. It introduces a new group of triggers that include geometric and kinematic selections on L1 objects. The low-$E_T$ trigger makes use of this L1 topological selection by accepting events where the largest energy deposit (and second-largest, if there is one) is required to have $E_T > 30$ GeV deposited in the HCal, with the additional condition that there are no energy deposits in the ECal with $E_T > 3$ GeV within a cone of size $\Delta R = 0.2$ around the HCal energy deposit. This veto on ECal deposits ensures a high value of $E_H/E_{EM}$ at L1, rejecting a large portion of background events. The trigger rate obtained with this condition is low enough to allow the $E_T$ threshold to be kept as low as 30 GeV. This looser $E_T$ requirement increases the efficiency for the low-$m_\phi$ signal models (those with $m_\phi \leq 200$ GeV).

In the high-level trigger (HLT), the selection algorithm for the CalRatio triggers is the same regardless of the L1 selection. Calorimeter deposits are clustered into jets using the anti-$k_t$ algorithm [61] with radius parameter $R = 0.4$. The standard jet cleaning requirements [62] applied in most ATLAS analyses reject jets with high values of $E_H/E_{EM}$, one of the main characteristics of the displaced hadronic jets, and are therefore not included in these triggers. A dedicated cleaning algorithm for jets created in the HCal (referred to as CalRatio jet cleaning) is applied instead, with no requirements on the jet $E_H/E_{EM}$. At least one of the HLT jets passing the CalRatio jet cleaning is required to satisfy $E_T > 30$ GeV, $|\eta| < 2.5$ and $\log_{10}(E_H/E_{EM}) > 1.2$. Jets satisfying these requirements are used to determine $0.8 \times 0.8$ regions in $\Delta \eta \times \Delta \phi$ centred on the jet axis in which to perform tracking. Triggering jets are required to have no tracks with $p_T > 2$ GeV within $\Delta R = 0.2$ of the jet axis. Finally, jets satisfying all of the above criteria are required to pass a BIB removal algorithm that relies on cell timing and position. Muons from BIB enter the HCal horizontally and may radiate a photon via bremsstrahlung, generating an energy deposit that may be reconstructed as a signal-like jet. Deposits due to BIB are expected to have a very specific time distribution [63]. The algorithm identifies events as containing BIB if the triggering jet has at least four HCal-barrel cells at the same $\phi$ and in the same calorimeter layer with timing consistent with that of a BIB deposit. In both CalRatio triggers, events identified as BIB by the BIB algorithm are saved in the BIB dataset and events with no triggering jets identified as BIB are saved in the main dataset.

The trigger is also active in so-called empty bunch crossings. These are crossings where protons are absent in both beams and isolated from filled bunches by at least five unfilled bunches on either side. Events in empty bunch crossings that have at least one $0.2 \times 0.2$ ($\Delta \eta \times \Delta \phi$) calorimeter energy deposit with $E_T > 30$ GeV at L1, and which pass the HLT selection algorithm, are stored in the cosmic-ray dataset.

The trigger efficiency for simulated signal events is defined as the fraction of jets spatially matched to one of the generated LLPs (hereafter called truth LLPs) that fire the trigger. The trigger efficiency as a function of triggering LLP particle-level $p_T$ is shown in Fig. 1 (left) for two signal samples. Only LLPs decaying in the HCal are considered in this plot. The high-$E_T$ CalRatio trigger, which is seeded by the high-$E_T$ L1 trigger, starts to be efficient for LLPs with $p_T > 100$ GeV and reaches its plateau at 150–200 GeV. The low-$E_T$ CalRatio trigger (seeded by the low-$E_T$ L1 trigger) recovers efficiency for a large portion of the LLPs with $p_T < 100$ GeV. The main source of efficiency loss in these triggers comes from the track isolation, followed by the combination of requirements on jet $E_T$ and $E_H/E_{EM}$. Fig. 1 (right) shows the LLP $p_T$ distribution for all the signal samples considered in the analysis. The combination of these figures shows how the high-$E_T$ CalRatio trigger gives a higher efficiency for models with $m_\phi > 200$ GeV, where the LLP $p_T$ distributions peak between 150 and 500 GeV. For signal models with $m_\phi$ up to 200 GeV, the LLP $p_T$ distributions peak between 30 and 100 GeV and hence the low-$E_T$ CalRatio trigger performs better. Thus, low-$m_\phi$ models are searched for using the low-$E_T$ dataset: despite the reduced integrated luminosity, a higher sensitivity is obtained than if
the high-$E_T$ dataset had been used. Conversely, models with $m_{\Phi} > 200$ GeV are studied using the high-$E_T$ dataset.

The trigger efficiency also depends strongly on the LLP decay position, as shown for three samples of simulated signal events in Fig. 2. The efficiency as a function of LLP decay length in the $x$-$y$ plane is shown for LLPs decaying in the barrel ($|\eta| < 1.4$); the efficiency as a function of the decay position in the $z$-direction is shown for LLPs decaying in the HCal endcaps (1.4 $\leq |\eta| < 2.5$). The selection is most efficient in the HCal for both triggers.

Events used in the analysis are required to pass the trigger requirements and contain a primary vertex (PV) with at least two tracks with $p_T > 400$ MeV. Tracks used in the jet and event selection hereafter are required to pass the track selection: they must originate from the PV and have $p_T > 2$ GeV.

The jets used in this analysis are selected by applying the following quality selections: $p_T > 40$ GeV, $|\eta| < 2.5$, pass CalRatio jet cleaning. These jets are referred to as clean. To select events with trackless jets, an additional event-level variable, $\sum \Delta R_{\text{min}}(\text{jet, tracks})$, is used. The quantity $\Delta R_{\text{min}}(\text{jet, tracks})$ is defined as the angular distance between the jet axis and the closest track with $p_T > 2$ GeV, and $\sum \Delta R_{\text{min}}(\text{jet, tracks})$ is calculated by summing this distance over all the clean jets with $p_T > 50$ GeV. Events with
no displaced decays have a very small value of this variable. Every displaced jet contributing to the sum causes a considerable increase in the value, making this variable a good discriminator between signal and multijet background. For an event to pass the analysis preselection, it is required to have passed the trigger, to contain at least two clean jets and to have \( \sum \Delta R_{\text{min}}(\text{jet, tracks}) > 0.5 \). After preselection, \( \sum \Delta R_{\text{min}}(\text{jet, tracks}) \) still has good discrimination power and it is used in the data-driven background estimation described in Sect. 5.

4.1 Displaced jet identification

Each clean jet is evaluated by a multilayer perceptron (MLP) (implemented in the Toolkit for Multivariate Data Analysis [64]) to predict the radial and longitudinal decay positions \((L_{xy} \text{ and } L_z)\) of the particle that produced the jet, using the jet’s fraction of energy deposited in each of the ECals and HCals as input variables. The MLP was trained on simulated signal samples with \( m_\Phi \) in the range [200, 1000] GeV, using only jets matched to a truth LLP. No requirements at event level (trigger and preselection) were applied in order to have as large a data sample as possible. In addition, avoiding the preselection allows the MLP to identify the decay position of prompt jets, which is useful when applied to SM jets. The MLP training procedure took as input the truth-level \( L_{xy} \) and \( L_z \) decay positions of the LLP as well as the fraction of the jet energy in each calorimeter layer, and finally the jet’s direction in \( \eta \).

The left-hand plot of Fig. 3 compares \( L_{xy} \) of a truth LLP against the MLP prediction. It shows clearly the different calorimeter layers, since decays in the same layer lead to constant MLP radial decay position prediction even as the truth decay position changes. However, the overall prediction in \( L_{xy} \) aligns closely with the truth decay position. The right plot shows the longitudinal decay position, \( L_z \). It shows a clear correlation between prediction and truth for the whole range of the forward calorimeter with less obvious layering, since the LLP direction of travel in the endcaps is more oblique with respect to the calorimeter layers than in the barrel. The radial and longitudinal decay positions predicted by the MLP are useful discriminators between signal jets from LLP decays in the calorimeters and prompt jets from SM backgrounds.

The per-jet BDT is used to separate jets into three classes: signal-like jets, SM multijet-like jets and BIB-like jets. With that purpose, it is trained using three samples. The signal sample contains jets from signal events for a range of models with \( m_\Phi \) in the range 125 – 1000 GeV, where only jets matched to LLPs decaying outside the ID (with \( L_{xy} > 1250 \text{ mm} \) if they decay in the barrel or \( L_z > 3500 \text{ mm} \) if they decay in the endcaps) are considered. The SM multijet training sample consists of jets from the simulated multijet events described in Sect. 3.2. Finally, the BIB sample is made of jets from the BIB dataset, where only the triggering jet in each event is used. The triggering jet is identified as BIB by the trigger BIB algorithm: the event contains a line of at least four HCal-barrel cells in the same \( \phi \) as the triggering jet, consistent with BIB timing. Hence, the triggering jet corresponds to a BIB jet in most cases, which is confirmed by the \( \phi \) and \( z \) vs. time plots showing the typical shapes of BIB. Using only the triggering jet reduces the risk of contamination from multijet events. In all cases, only clean jets are considered.

The per-jet BDT inputs are the MLP \( L_{xy} \) and \( L_z \) predictions, track variables, and jet properties. The track variables include the sum of \( p_T \) of all tracks passing track selection within \( \Delta R = 0.2 \) of the jet axis, and the maximum \( p_T \) of such tracks. The jet properties are: the radius, shower centroid, energy density and fraction of energy in first HCal layer of the cluster with the highest \( p_T \); the longitudinal and transverse distance from this cluster to the jet shower center; jet
Fig. 4 The distributions of the per-jet BDT weights for a multijet sample, a BIB sample and five signal samples. For the signal samples, the weights for clean jets matched to an LLP decaying in the calorimeter are shown. The multijet and BIB distributions include weights for all clean jets in the event.

\[ p_T; \text{ and the compatibility of the jet timing with the expected timing of a BIB deposit.} \]

The jet \( p_T \) spectrum is very different in each of the three training samples, and therefore jets in each sample are weighted such that the jet \( p_T \) distribution is flat. The weighting is done independently in each training sample. Since the jet \( p_T \) is correlated with a number of BDT input variables, the jet \( p_T \) is also included as a variable in the BDT.

The output of the per-jet BDT is a set of three weights that sum to unity: signal-weight, BIB-weight and multijet-weight, shown in Fig. 4. The signal-weight distribution provides a clear separation between signal jets and both types of background jets. The BIB-weight distributions for signal and multijet jets peak at intermediate values. Jets from the BIB sample with low BIB-weight scores (< 0.34) display SM multijet-like qualities and are likely to result from SM jet contamination in the BIB sample. Jets with higher BIB-weight values display the expected timing behaviour of particles originating from BIB. The per-jet BDT is able to separate these with some precision, assigning values between 0.34 and 0.35 to BIB particles crossing the detector through the innermost layer of the HCal and higher values (> 0.35) to BIB in outer HCal layers. The per-jet BDT has better signal-to-background discrimination for high-\( m_\Phi \) models than for low-\( m_\Phi \) models. The main reason for this lies in the \( p_T \) distribution (see Fig. 1). Both the BIB and pile-up jets have relatively soft \( p_T \), and even though these backgrounds are mitigated by the jet-cleaning requirements, their remaining contributions are harder to distinguish at low \( p_T \). The presence of pile-up jets has two effects: on the one hand, they can leave energy deposits in the ECal, changing the fraction of energy per calorimeter layer and worsening the signal-to-background discrimination. On the other hand, pile-up jets’ tracks do not point back to the PV in many cases and hence are not considered for track isolation. These jets can be reconstructed as nearly trackless, making them more similar to signal.

4.2 Event selection

A per-event BDT is defined with the main objective of discriminating BIB events from signal events. A combination of signal samples is used as signal in the training while the BIB dataset events are used as background.

The two jets with the highest per-jet signal-weight in the event (\( \text{CalRatio jet candidates} \)) and the two jets with the highest per-jet BIB-weight in the event (\( \text{BIB jet candidates} \)) are selected and their per-jet weights are used as input variables to the per-event BDT. Other event-level variables such as \( H_T^{\text{miss}} / H_T \), where \( H_T \) is the scalar sum of jet transverse
moments and $H_T^{\text{miss}}$ is the magnitude of the vectorial sum of transverse moments of these jets, and the distance $\Delta R$ between the two CalRatio jet candidates are used in the training.

As mentioned in the previous subsection, signal jets with low $p_T$ are harder to discriminate from background. For this reason, and to obtain an optimal signal-to-background discrimination at all $p_T$, two versions of the per-event BDT are trained: one for the analysis of the high-$E_T$ dataset, and another for the low-$E_T$ dataset. They only differ in the signal samples used for training and in the triggers required to select events. The high-$E_T$ per-event BDT training uses a combination of low-, intermediate- and high-mass signal samples in events passing the high-$E_T$ CalRatio trigger. The low-$E_T$ per-event BDT training uses a combination of low-$m_\Phi$ signal samples and only events passing the low-$E_T$ CalRatio trigger. Figure 5 shows the distribution of the per-event BDTs from five signal samples, as well as from the main data and BIB data. The BIB training sample contains SM multijet events in addition to the BIB jet that caused them to be selected by the trigger. Consequently, even if no multijet sample is used in the training, the per-event BDT is able to discriminate signal from BIB as well as from multijet background. This can be seen in Fig. 5 by comparing the BDT results in the main data and the BIB datasets, especially in the low-$E_T$ per-event BDT output. Using time and $z$-coordinate measurements, it has been checked that events with low per-event BDT values ($< -0.2$) have the typical characteristics of BIB, while events with intermediate values (between $-0.2$ and $0.2$) are multijet-like.

The simulated distributions of the variables used as BDT inputs (for both the per-jet and per-event BDTs) are compared with data, and good agreement is generally observed. The small remaining discrepancies are propagated into an uncertainty in the modelling of BDT input variables, which is described in Sect. 6.

Two selections are defined, referred to as the high-$E_T$ selection and the low-$E_T$ selection, which are optimised to give maximum sensitivity for high-$m_\Phi$ models and low-$m_\Phi$ models, respectively.

Event cleaning selections are applied to remove as much BIB background as possible: trigger matching (at least one of the CalRatio jet candidates has to be matched to the jet that fired the trigger), and a timing window of $-3 < t < 15$ ns for the CalRatio jet candidates and for the BIB jet candidates. Furthermore, the per-event BDT output is required to satisfy high-$E_T$ per-event BDT $> 0.1$ and low-$E_T$ per-event BDT $> 0.1$ in the high-$E_T$ and low-$E_T$ selections, respectively. These requirements ensure that the only source of background contributing to the final selection is multijet events.

The final selection is optimised to maximise the signal-to-background ratio in each search region. Variables with good signal-to-background discrimination at event level are used, such as $H_T^{\text{miss}}$/$H_T$ and $\sum_{j_1,j_2} \log_{10}(E_{T_{\text{had}}}/E_{T_{\text{EM}}})$, where $j_1$ and $j_2$ refer to the CalRatio jet candidates. The quantity $H_T^{\text{miss}}$/$H_T$ has a value close to 1 for BIB events, but it has a softer distribution for signal. This variable replaces the $E_T^{\text{miss}} < 30$ GeV requirement applied in the 8 TeV analysis [37] (where $E_T^{\text{miss}}$ is the magnitude of the negative vector transverse momentum sum of the reconstructed and calibrated physics objects), which was very useful for reducing the multijet background with only a small effect on the efficiency of low-$m_\Phi$ models. However, it significantly lowered the efficiency for the high-$m_\Phi$ models due to larger portions of the high-$p_T$ jets escaping the calorimeters (punch-through), generating fake $E_T^{\text{miss}}$. The elimination of this requirement improves the sensitivity of the analysis to the high-$m_\Phi$ models by a large factor, while the improvement is less noticeable for low-$m_\Phi$ models. The following additional requirements are applied for the high-$E_T$ selection: $\sum_{j_1,j_2} \log_{10}(E_{T_{\text{had}}}/E_{T_{\text{EM}}}) > 1$, $p_T(j_1) > 160$ GeV, $p_T(j_2) > 100$ GeV, and $H_T^{\text{miss}}$/$H_T < 0.6$. The low-$E_T$ selection requires $\sum_{j_1,j_2} \log_{10}(E_{T_{\text{had}}}/E_{T_{\text{EM}}}) > 2.5$, $p_T(j_1) > 80$ GeV, and $p_T(j_2) > 60$ GeV.
5 Background estimation

The data-driven ABCD method is used to estimate the contribution from the dominant background (SM multijet events) to the final selection. The standard ABCD method relies on the assumption that the distribution of background events can be factorised in the plane of two relatively uncorrelated variables. In this plane, the method uses three control regions (B, C and D) to estimate the contribution of background events in the search region (A). If all the signal events are concentrated in region A, the number of background events in region A can be predicted from the population of the other three regions using 

\[ N_A = (N_B \cdot N_C)/N_D, \]

where \( N_X \) is the number of background events in region \( X \). In reality, some signal events may lie outside of region A. A modified ABCD method is used to account for non-zero signal contamination in regions B, C and D. The modified ABCD method involves fitting to background and signal models simultaneously. The background component of the yields in regions A, B, C and D are constrained to obey the standard ABCD relation, within the bounds of the ABCD method uncertainty (described below).

The validity of the ABCD method is tested by applying it to two validation regions (VRs). These are similar to the main selections, but have modified requirements and boundaries for the ABCD plane variables, to ensure orthogonality to the high-\( E_T \) and low-\( E_T \) selections. The VR for the high-\( E_T \) selection (VR\(_{\text{high-}E_T}\)) is defined as the nominal selection except for requiring 100 < \( p_T(j_1) < 160 \) GeV and it is evaluated in the ABCD plane defined within 0.1 < high-\( E_T \) per-event BDT < 0.22. The VR for the low-\( E_T \) selection (VR\(_{\text{low-}E_T}\)) is defined as the nominal selection and it is evaluated in the ABCD plane defined within 0.1 < low-\( E_T \) per-event BDT < 0.22.

In both VRs, the correlation observed between the two variables defining the ABCD plane is negligible (< 3% in main data) and signal contamination in region A is small. In all cases, the estimated number of background events is in good agreement with the number of data events observed in region A, as summarised in Table 2.

The uncertainty in the data-driven background estimate is studied using a dijet-enriched sample. This sample is selected using a single-jet-based trigger and vetoing on the CalRatio triggers to make sure that the event selection is orthogonal to the one used in the main analysis. The ABCD planes are then defined similarly to those in the main analysis, but adjusting the boundaries in regions A, B, C and D to reduce the effect of statistical fluctuations in the estimation of the number of events passing each stage of the analysis selections. For both the high-\( E_T \) and low-\( E_T \) selections, the number of BIB events satisfying all selection criteria is well within the uncertainty in the number of events passing all selections in the main dataset. Furthermore, the events from the BIB dataset that pass the selection were checked, and found to display properties of multijet events. In particular, their \( \phi \) and \( z \) vs time distributions do not show the typical shape of BIB. The events from the main dataset that pass the event cleaning were also checked, and were found not to display the properties of BIB.

The second check is to ensure that almost all the cosmic-ray background is removed, using the cosmic-ray dataset. The estimated number of events passing each stage of the selection is listed in Table 1 for the high-\( E_T \) (low-\( E_T \)) selection. In both cases the number is also within the statistical uncertainty for the number of events entering the selection in the main dataset.
dijet events in region A given by the method. The difference between the estimated and observed numbers of events in region A is taken as the systematic uncertainty associated with the method: 22% in the high-$E_T$ ABCD plane and 25% in the low-$E_T$ plane. The size of the statistical component of these uncertainties is 17% and 20%, respectively.

The yields in each region of the main high-$E_T$ and low-$E_T$ selections are shown in the Table 3 alongside the final background estimate calculated from a simultaneous background-only fit to all regions using the statistical model described in Sect. 7. The expected background in each region is allowed to float so long as the ABCD relation is satisfied, with a Poisson constraint on the observed number of events in the corresponding region. If the observed data in region A are ignored in the fit by removing the Poisson constraint on region A, the background estimate is the same as that expected from the ABCD relation ($N_{\text{bkg}}^\text{A} = (N_{\text{bkg}}^\text{B} \cdot N_{\text{bkg}}^\text{C}) / N_{\text{bkg}}^\text{D}$), but with all sources of uncertainty accounted for. This corresponds to the a priori (pre-unblinding) background estimate. The a posteriori (post-unblinding) background estimate, which is used for the purposes of statistical interpretation, is obtained from the same background-only simultaneous fit to all regions, taking the observed number of events in A into account. Here also the ABCD relation is imposed, within the uncertainty of the ABCD method. When performing a signal-plus-background fit during the statistical interpretation, the estimated background can vary as a function of the signal strength.

6 Systematic uncertainties

The uncertainty in the data-driven ABCD method for the background estimate is discussed in Sect. 5, and found to be 22% in the high-$E_T$ ABCD plane and 25% in the low-$E_T$ plane.

Several uncertainties related to modelling, theory and reconstruction affect the estimated signal yield. The jet-energy scale and jet-energy resolution introduce uncertainties in the signal yield of 1–9% and 1–5%, respectively, depending on the model, where the high-$m_\phi$ models are least affected. These uncertainties are calculated using the
The uncertainty in the signal trigger efficiency is estimated by studying how well modelled the three main HLT variables (jet $E_T$ and $\log_{10}(E_H/E_{EM})$, and $p_T$ of tracks within the jet) are between HLT- and offline-reconstructed quantities in data and Monte Carlo (MC) simulation. A tag-and-probe technique using standard jet triggers is used to obtain a pure sample of multijet events in both data and MC simulation. Scale factors are derived that represent the degree of mis-modelling in each variable, and are applied in an emulation of the CalRatio triggers. The change in yield relative to the nominal (unscaled) trigger emulation after the full analysis selection is taken as the size of the systematic uncertainty, which is 2% or less for all models.

Events in MC simulation are reweighted to obtain the correct pileup distribution. A variation in the pileup reweighting of MC is included to cover the uncertainty on the ratio between the predicted and measured inelastic cross-section in the fiducial volume defined by $M_X > 13$ GeV where $M_X$ is the mass of the hadronic system [66]. The uncertainty in the pile-up reweighting of the reconstructed events in the MC simulation is estimated by comparing the distribu-

### Table 1: Sequential impact of each requirement on the number of events

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Main data</th>
<th>BIB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td>Pass trigger, 2 clean jets and $\sum \Delta R_{\text{min}} &gt; 0.5$</td>
<td>2180349</td>
</tr>
<tr>
<td>Event cleaning</td>
<td>Low-$E_T$ per-event BDT &gt; 0.1</td>
<td>319.1</td>
</tr>
<tr>
<td>High-$E_T$ selection</td>
<td>$\sum_{j1,j2} \log_{10}(E_H/E_{EM}) &gt; 1$</td>
<td>6.26</td>
</tr>
<tr>
<td>Region A</td>
<td>$p_T(j1) &gt; 80$ GeV</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>$p_T(j2) &gt; 60$ GeV</td>
<td>3.5</td>
</tr>
</tbody>
</table>

### Event cleaning

- $H_T^\text{miss}/H_T < 0.6$
- $p_T(j1) > 160$ GeV
- $p_T(j2) > 100$ GeV

### High-$E_T$ selection

- $E_T > 15$ ns
- $3 < t < 15$ ns
- $\sum_{j1,j2} \log_{10}(E_H/E_{EM}) > 2.5$
- $p_T(j1) > 80$ GeV
- $p_T(j2) > 60$ GeV

### Region A

- $\sum \Delta R_{\text{min}} > 0.5$
- $3 < t < 15$ ns
- $\sum_{j1,j2} \log_{10}(E_H/E_{EM}) > 2.5$
- $p_T(j1) > 80$ GeV
- $p_T(j2) > 60$ GeV

### Preselection

- Pass trigger, 2 clean jets and $\sum \Delta R_{\text{min}} > 0.5$
- $E_T > 15$ ns
- $3 < t < 15$ ns
- $\sum_{j1,j2} \log_{10}(E_H/E_{EM}) > 2.5$
- $p_T(j1) > 80$ GeV
- $p_T(j2) > 60$ GeV

### Region A

- $\sum \Delta R_{\text{min}} > 0.5$
- $3 < t < 15$ ns
- $\sum_{j1,j2} \log_{10}(E_H/E_{EM}) > 2.5$
- $p_T(j1) > 80$ GeV
- $p_T(j2) > 60$ GeV

---

The quoted number for BIB corresponds to the number of events in the BIB dataset passing the selection. The number of events for the cosmic rays is estimated from the cosmic-ray dataset by applying corrections for the proportion of empty crossings relative to filled crossings, and a factor to account for the change in analysis efficiency in a zero-pile-up collision. Signal yields are quoted as a percentage of the total number of generated events.
validation of the number of primary vertices in the MC simulation with the one in data as a function of the instantaneous luminosity. Differences between these distributions are adjusted by scaling the mean number of $pp$ interactions per bunch crossing in the MC simulation and the $\pm 1\sigma$ uncertainties are assigned to these scaling factors. The effect on the signal event yields varies between 1 and 12% depending on the model. The low-$m_\Phi$ models are the most affected by this uncertainty.

The NNPDF2.3LO [51] PDF set was used when generating the signal samples. In addition to the nominal PDF, 100 PDF variations are also included in the set. The PDF uncertainty is evaluated by taking the standard deviation of signal event yield when each of these PDF variations is used instead of the nominal. The effect on the signal yield is between 3% and 8% depending on the signal sample, where the size of the uncertainty grows with $m_\Phi$.

A systematic uncertainty is included to account for potential mis-modelling of BDT input variables, using the same control sample of dijet events defined for the evaluation of the systematic uncertainty in the data-driven background estimate. In this control sample, the distributions of the inputs and outputs of the per-jet and per-event BDTs were studied, and were found to agree fairly well between data and MC simulation. The residual differences are translated into a systematic uncertainty in the signal efficiency by randomly varying the input variables according to their uncertainty and re-evaluating the BDTs for each signal event. The value of the resulting uncertainty is up to 2% depending on the model, where the largest uncertainties are assigned to the lower-$m_\Phi$ models.

Finally, the uncertainty in the integrated luminosity is around 2%. It is derived, following a methodology similar to that detailed in Ref. [67], and using the LUCID-2 detector for the baseline luminosity measurements [68], from calibration of the luminosity scale using $x-y$ beam-separation scans. This uncertainty affects all models equally.

### 7 Statistical interpretation

#### 7.1 Extraction of limits

A data-driven background estimation and signal hypothesis test is performed simultaneously in all regions. An overall profile likelihood function is constructed from the product of the Poisson probabilities of observing the number of events $N_{\text{obs}}^X$, given an expectation $N_{\text{exp}}^X$, in each region $X$, where $X = A, B, C, D$. The value of $N_{\text{exp}}^X$ in each region is the sum of: the expected signal yield $N_{\text{sig}}^X$, given by the number of simulated signal events entering region $X$ multiplied by the signal strength $\mu$ (the parameter of interest); and the expected background yield $N_{\text{bkg}}^X$. In the fit, the expected background yields are constrained to obey the ABCD relation $N_{\text{exp}}^A = (N_{\text{bkg}}^B \times N_{\text{bkg}}^C) / N_{\text{bkg}}^D$. This reduces the number of degrees of freedom of the fit by one as $N_{\text{obs}}^A = m N_{\text{bkg}}^B + \mu N_{\text{sig}}^A$ and $N_{\text{obs}}^C = m N_{\text{bkg}}^D + \mu N_{\text{sig}}^C$, where $m$ is a free parameter. Since the Poisson constraints only apply to $N_{\text{obs}}^X$ relative to $N_{\text{exp}}^X$, it follows that the background prediction may change dynamically in the fit as a function of the signal strength.

As can be seen in Table 3, no excess of events is observed in region A for either of the analysis selections. The CLs method [69] is therefore used to set upper limits on $\sigma(\Phi) \times B_{\Phi \rightarrow \mu \mu}$ in the benchmark HS model.

Systematic uncertainties for signal, background and luminosity are represented by nuisance parameters. Each nuisance parameter is assigned a Gaussian constraint of relevant width (see Sect. 6). An asymptotic approach [70] is used to compute the CLs value, and the limits are defined by the region excluded at 95% confidence level (CL). The asymptotic approximation was tested and found to give consistent results with limits obtained from ensemble tests.

Since each signal sample was generated for a particular LLP proper decay length, it is necessary to extrapolate the signal efficiency to other decay lengths to obtain limits as a function of $\tau$. This is achieved by using a weighting method, which is applied separately to each signal sample. The weight to be assigned to a displaced jet with lifetime $\tau_{\text{new}}$ is obtained from the sample generated with lifetime $\tau_{\text{gen}}$ by:

$$w(t) = \frac{\tau_{\text{gen}}}{\tau_{\text{new}}} \cdot \frac{\exp(-t/\tau_{\text{gen}})}{\exp(-t/\tau_{\text{new}})}.$$ 

The quantity $t$ is the proper decay time of the LLP that gives rise to the displaced jet. In the benchmark HS model, the LLPs are pair produced, so each event is weighted by the product of the individual LLP weights. The weighted sample is used to evaluate the signal efficiency for $c_\tau \tau_{\text{new}}$. 

---

Table 2 Application of the ABCD method to the final high-$E_T$ and low-$E_T$ VRs. The column labelled “Estim. A” gives the estimated contribution of background in the search region A assuming no signal, as calculated using the ABCD method. The statistical uncertainty of this calculation is also given. A, B, C and D show the number of observed events in each region. Only statistical uncertainties are considered in this table.

<table>
<thead>
<tr>
<th>Validation selections</th>
<th>Estim. A</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR$_{\text{high}-E_T}$</td>
<td>66 ± 15</td>
<td>70</td>
<td>64</td>
<td>57</td>
<td>55</td>
</tr>
<tr>
<td>VR$_{\text{low}-E_T}$</td>
<td>54 ± 17</td>
<td>36</td>
<td>35</td>
<td>34</td>
<td>22</td>
</tr>
</tbody>
</table>
### Table 3 Application of the modified ABCD method to the final high-\(E_T\) and low-\(E_T\) selections. The columns A, B, C and D contain the number of observed events in each region. “Estim. A” refers to the estimated contribution of background in the search region A assuming no signal, as calculated using the full statistical model described in Sect. 7. The a priori estimate refers to the “pre-unblinding” case, where the data in region A are ignored by removing the Poisson constraint in that region and the signal strength is fixed to zero. This corresponds to the simple \(N_{bkg}^{A}= (A_B^{bkg} \cdot A_C^{bkg})/A_B^{bkg}\) relation. The a posteriori estimate refers to the “post-unblinding” case, including the observed data in region A into the background only global fit, obtained by fixing the signal strength to 0. Only the fitted value in region A is shown, though the fitted number of events in regions B, C and D fluctuate as well. All experimental and statistical uncertainties are included in the quoted error.

<table>
<thead>
<tr>
<th>Main selections</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Estim. A (a priori)</th>
<th>A</th>
<th>Estim. A (a posteriori)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-(E_T) selection</td>
<td>9</td>
<td>187</td>
<td>253</td>
<td>6.7 (_{+3.2}^{-2.3})</td>
<td>10</td>
<td>8.5 (_{+2.3}^{-2.0})</td>
</tr>
<tr>
<td>Low-(E_T) selection</td>
<td>2</td>
<td>70</td>
<td>57</td>
<td>2.5 (_{+2.5}^{-1.4})</td>
<td>7</td>
<td>5.3 (_{+2.1}^{-1.6})</td>
</tr>
</tbody>
</table>

#### Fig. 7 The extrapolated signal efficiencies as a function of proper decay length of the \(s\) for several simulated samples in the low-\(E_T\) (left) and high-\(E_T\) (right) selections. The vertical error bars represent the statistical uncertainties.

The upper limit at a given \(c \tau\) is then obtained by scaling the limit at \(c \tau_{\text{gen}}\) by the ratio of signal efficiencies at \(c \tau\) and \(c \tau_{\text{gen}}\). This procedure for extrapolating the efficiency to different lifetimes was checked by comparing the extrapolated efficiency derived from the main simulated samples with the measured efficiency of samples with alternative LLP lifetime assumptions. These were found to agree within statistical uncertainties. Figure 7 shows the extrapolated efficiency for the signal samples with \(m_{\Phi}\) of 125 and 200 GeV with the low-\(E_T\) selection applied, alongside the efficiency for signal samples with \(m_{\Phi}\) of 400 GeV, 600 GeV, and 1 TeV signal samples with the high-\(E_T\) selection applied.

The observed and expected limits for two example signal models can be seen in Fig. 8. The observed limits for all considered models are summarised in Fig. 9. The expected limits correspond to those obtained using the a posteriori background estimate, which is given in Table 3. This explains why the observed and expected limits may appear closer than anticipated from the observed and expected numbers of events in region A using the simple ABCD relation.

For a mediator similar to the Higgs boson and of mass \(m_{\Phi}\) = 125 GeV, the limits are presented divided by the SM Higgs boson gluon–gluon fusion production cross section for \(m_H = 125\) GeV, assumed to be 48.58 pb at 13 TeV [71]. For such models, decays of neutral scalars with masses between 5 and 55 GeV are excluded for proper decay lengths between 5 cm and 5 m depending on the LLP mass (assuming a 10% branching ratio). Compared with the 8 TeV results, the limits for models with \(m_{\Phi} = 125\) GeV are typically a factor 10 more stringent around 20 cm and a factor 10 less stringent around 50 m.

For \(m_{\Phi} = 200\) GeV, cross section times branching ratio values above 1 pb are ruled out between 5 cm and 7 m depending on the scalar mass. For models with \(m_{\Phi} = 400\) GeV, \(m_{\Phi} = 600\) GeV, and \(m_{\Phi} = 1000\) GeV, \(\sigma (\Phi) \times B_{\Phi \rightarrow ss}\) values above 0.1 pb are ruled out at 95% CL between about 12 cm and 9 m, 7 cm and 20 m, and 4 cm and 35 m respectively, depending on the scalar masses. The limits are significantly more stringent than the 8 TeV results across the whole lifetime range, and in some cases limits are set on combinations of \(m_{\Phi}\) and \(m_s\) that were not previously studied.

### 7.2 Combination of results with MS displaced jets search

In this section the limits derived in Sect. 7.1 are combined with the results for the comparable models from
A search for pair-produced long-lived particles decaying in the ATLAS calorimeter is presented, using data collected during $pp$ collisions at the LHC in 2016, at centre-of-mass energy of 13 TeV. The dataset size is 10.8 fb$^{-1}$ or 33.0 fb$^{-1}$ depending on whether the data were collected using a low- or high-$E_T$ dedicated trigger. Benchmark hidden-sector models are used to set limits, where the mediator’s mass ranges between 125 and 1000 GeV, while the long-lived scalar’s mass range between 5 and 400 GeV. The search selects events with two signal-like jets (which are typically narrow, trackless, and with a large fraction of their energy in the hadronic calorimeter) using machine-learning tech-

The muon spectrometer (MS) displaced-jets analysis [40]. The MS analysis searches for neutral LLPs decaying at the outer edge of the HCal or in the MS. These decays result in secondary-decay vertices that can be reconstructed as displaced vertices in the MS. The analysis considers events containing either two displaced vertices in the MS or one displaced vertex together with prompt jets or $E_T^{\text{miss}}$. Some of the benchmark models used in the MS vertex search are the same models considered in the search described in this paper. Therefore a combination of the results of these two complementary analyses can be performed.

The orthogonality of the CalRatio (CR) and MS analyses was checked in both data and simulated signal to ensure the final selections were statistically independent. The combination is performed using a simultaneous fit of the likelihood functions of each analysis. The signal strength as well as the nuisance parameter for the luminosity uncertainties is chosen to be the same for the CR and MS likelihoods. The signal uncertainties are chosen to be uncorrelated, since they are dominated by different experimental uncertainties in the two searches. The effect of correlating the signal uncertainties was studied by comparing the limit obtained with no correlation in signal uncertainties to that obtained with correlation of relevant signal uncertainties. The effect on the combined limits was found to be negligible. The background estimate in each analysis is data-driven and the two estimates are therefore not correlated.

As in the individual searches, the asymptotic approach is used to compute the CL$_S$ value, and the limits are defined by the region excluded at 95% CL. The limits are calculated using a global fit, where the overall likelihood function is the product of the individual likelihood functions of the searches to be combined. The limits are calculated separately at each point in the $cT$ range of interest, where in each case the signal efficiency is scaled by the result of the lifetime extrapolation.

The observed and expected limits for two example signal models are shown in Fig. 10. For the models with $m_\Phi = 125$ GeV, the MS analysis has higher sensitivity than the CR analysis at large decay lengths. For short decay lengths ($<10$ cm) the sensitivities of the two analyses are comparable and the combination of their limits provides a slight improvement. The limits for intermediate masses, $m_\Phi = 200$ and 400 GeV, show a clear complementarity of the analyses; the CR limits, which improve with $m_\Phi$, are stronger at shorter decay lengths, while the MS analysis sets stronger limits at large decay lengths. In this case the combination of the two analyses improves on the individual limits over the full range of decay lengths. For higher masses, $m_\Phi \geq 600$ GeV, the CR analysis is in general more sensitive than the MS analysis. Even in this case, the combination provides a modest improvement on the CR-only limit at long decay lengths.

8 Conclusion

The observed and expected limits for two example signal models are shown in Fig. 10. For the models with $m_\Phi = 125$ GeV, the MS analysis has higher sensitivity than the CR analysis at large decay lengths. For short decay lengths ($<10$ cm) the sensitivities of the two analyses are comparable and the combination of their limits provides a slight improvement. The limits for intermediate masses, $m_\Phi = 200$ and 400 GeV, show a clear complementarity of the analyses; the CR limits, which improve with $m_\Phi$, are stronger at shorter decay lengths, while the MS analysis sets stronger limits at large decay lengths. In this case the combination of the two analyses improves on the individual limits over the full range of decay lengths. For higher masses, $m_\Phi \geq 600$ GeV, the CR analysis is in general more sensitive than the MS analysis. Even in this case, the combination provides a modest improvement on the CR-only limit at long decay lengths.
Fig. 9 The observed limits for the $\Phi$ masses of 125, 200, 400, 600 and 1000 GeV

The background estimation is performed using the data-driven ABCD method. No significant excess is observed in either signal region. The CLs method is therefore used to set 95% CL limits on $\sigma(\Phi) \times B_{\Phi \rightarrow ss}$ as a function of LLP decay length. For a mediator similar to the Higgs boson and of mass $m_{\Phi} = 125$ GeV, decays of neutral scalars with masses between 5 and 55 GeV are excluded for proper decay lengths between 5 cm and 5 m depending on the LLP mass (assuming a 10% branching ratio). For $m_{\Phi} = 200$ GeV, cross section times branching ratio values above 1 pb are ruled out between 5 cm and 7 m depending on the scalar mass. For models with $m_{\Phi} = 400$ GeV, $m_{\Phi} = 600$ GeV, and $m_{\Phi} = 1000$ GeV, $\sigma(\Phi) \times B_{\Phi \rightarrow ss}$ values above 0.1 pb are ruled out between about 12 cm and 9 m, 7 cm and 20 m, and 4 cm and 35 m respectively, depending on the scalar masses. A combination of the limits with the results of a similar ATLAS search looking for displaced vertices in the muon spectrometer is performed. The resulting combined limits provide a summary of the ATLAS results for pair-produced neutral LLPs. The combined limits tend to follow the results from the most sen-
Fig. 10 Examples of the combined limits for models with $m_\Phi = 125$ GeV and $m_\Phi = 600$ GeV from the CR analysis and the MS analysis, which is separated into the MS 1-vertex plus $E_{\text{miss}}$ (MS1) and MS 2-vertex (MS2) components. The MS1 component of the MS displaced jet search was only applied to models with $m_\Phi = 125$ GeV. The expected limit is shown as a dashed line with shading for the $\pm 1\sigma$ band, while the observed is a solid line. The colours of the shading and solid and dashed lines refer to the limits from each analysis and their combination, as indicated in the legend.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMHE, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, The Netherlands; RCN, Norway; MINSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, UK; DOE and NSF, USA. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, Spain; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (The Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [72].

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk/). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/). This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413[opendata.cern.ch].]

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP³.

References


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>O. E. Vickey Boeriu</td>
<td>Department of Physics, University of Adelaide, Australia</td>
<td></td>
</tr>
<tr>
<td>G. H. A. Viehhauser</td>
<td>Physics Department, SUNY Albany, Albany, NY, USA</td>
<td></td>
</tr>
<tr>
<td>L. Vigan</td>
<td>Department of Physics, University of Alberta, Edmonton, CA</td>
<td></td>
</tr>
<tr>
<td>M. Villa</td>
<td>(a) Department of Physics, Ankara University, Ankara, Canada; (b) Istanbul Aydin University, Istanbul, Turkey</td>
<td></td>
</tr>
<tr>
<td>M. Villaplana Perez</td>
<td>(c) Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey</td>
<td></td>
</tr>
<tr>
<td>E. Vilucchi</td>
<td>LAPP, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy, France</td>
<td></td>
</tr>
<tr>
<td>M. G. Vincter</td>
<td>High Energy Physics Division, Argonne National Laboratory, Argonne, IL, USA</td>
<td></td>
</tr>
<tr>
<td>V. B. Vinogradov</td>
<td>Department of Physics, University of Arizona, Tucson, AZ, USA</td>
<td></td>
</tr>
<tr>
<td>A. Vishwakarma</td>
<td>Department of Physics, University of Texas at Arlington, Arlington, TX, USA</td>
<td></td>
</tr>
<tr>
<td>I. Vivarelli</td>
<td>Physics Department, National and Kapodistrian University of Athens, Athens, Greece</td>
<td></td>
</tr>
<tr>
<td>M. Vogel</td>
<td>Physics Department, National Technical University of Athens, Zografou, Greece</td>
<td></td>
</tr>
<tr>
<td>P. Vokac</td>
<td>Department of Physics, University of Austin, Texas, USA, USA</td>
<td></td>
</tr>
<tr>
<td>G. Volpi</td>
<td>(a) Bahcesehir University, Faculty of Engineering and Natural Sciences, Istanbul, USA; (b) Istanbul Bilgi University, Faculty of Engineering and Natural Sciences, Istanbul, USA; (c) Department of Physics, Bogazici University, Istanbul, USA; (d) Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey</td>
<td></td>
</tr>
<tr>
<td>E. Von Toerne</td>
<td>Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan</td>
<td></td>
</tr>
<tr>
<td>V. Vorobev</td>
<td>Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain</td>
<td></td>
</tr>
<tr>
<td>M. Vos</td>
<td>(a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b) Physics Department, Tsinghua University, Beijing, China; (c) Department of Physics, Nanjing University, Nanjing, China; (d) University of Chinese Academy of Science (UCAS), Beijing, China</td>
<td></td>
</tr>
<tr>
<td>J. H. Bossebeld</td>
<td>Institute of Physics, University of Belgrade, Belgrade, Serbia</td>
<td></td>
</tr>
<tr>
<td>Department of Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei, China</td>
<td>(b) Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao, China</td>
<td>(c) School of Physics and Astronomy, Shanghai Jiao Tong University, KLPPAC-MoE, SKLPPC, Shanghai, China</td>
</tr>
<tr>
<td>Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima, Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) Department of Physics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China</td>
<td>(b) Department of Physics, University of Hong Kong, Hong Kong, China</td>
<td>(c) Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China</td>
</tr>
<tr>
<td>Department of Physics, National Tsing Hua University, Hsinchu, Taiwan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Physics, Indiana University, Bloomington, IN, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy</td>
<td>(b) ICTP, Trieste, Italy</td>
<td>(c) Dipartimento Politecnico di Ingegneria e Architettura, Università di Udine, Udine, Italy</td>
</tr>
<tr>
<td>(a) INFN Sezione di Lecce, Zona Monte, Italy</td>
<td>(b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Milano, Milan, Italy</td>
<td>(b) Dipartimento di Fisica, Università di Milano, Milano, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Napoli, Naples, Italy</td>
<td>(b) Dipartimento di Fisica, Università di Napoli, Napoli, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Pavia, Pavia, Italy</td>
<td>(b) Dipartimento di Fisica, Università di Pavia, Pavia, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Pisa, Pisa, Italy</td>
<td>(b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Roma, Rome, Italy</td>
<td>(b) Dipartimento di Fisica, Sapienza Università di Roma, Rome, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Roma Tor Vergata, Rome, Italy</td>
<td>(b) Dipartimento di Fisica, Università di Roma Tor Vergata, Rome, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN Sezione di Roma Tre, Rome, Italy</td>
<td>(b) Dipartimento di Matematica e Fisica, Università Roma Tre, Rome, Italy</td>
<td></td>
</tr>
<tr>
<td>(a) INFN-TIFPA, Rome, Italy</td>
<td>(b) Università degli Studi di Trento, Trento, Italy</td>
<td></td>
</tr>
<tr>
<td>Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Iowa, Iowa City, IA, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Physics and Astronomy, Iowa State University, Ames, IA, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint Institute for Nuclear Research, Dubna, Russia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Brazil</td>
<td>(b) Universidade Federal do Rio de Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil</td>
<td>(c) Universidade Federal de São João del Rei (UFSJ), São João del Rei, Brazil</td>
</tr>
<tr>
<td>KEK, High Energy Accelerator Research Organization, Tsukuba, Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate School of Science, Kobe University, Kobe, Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Krakow, Poland</td>
<td>(b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland</td>
<td></td>
</tr>
<tr>
<td>Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty of Science, Kyoto University, Kyoto, Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyoto University of Education, Kyoto, Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka, Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physics Department, Lancaster University, Lancaster, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana, Slovenia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School of Physics and Astronomy, Queen Mary University of London, London, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Physics, Royal Holloway University of London, Egham, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Physics and Astronomy, University College London, London, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Louisiana Tech University, Ruston, LA, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fysiska institutionen, Lunds universitet, Lund, Sweden</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>