Constraining electroweak and strongly charged long-lived particles with CheckMATE

Nishita Desai\textsuperscript{1,a}, Florian Domingo\textsuperscript{2,b}, Jong Soo Kim\textsuperscript{3,c}, Roberto Ruiz de Austri Bazan\textsuperscript{4,d}, Krzysztof Rolbiecki\textsuperscript{5,e}, Mangesh Sonawane\textsuperscript{6,f}, Zeren Simon Wang\textsuperscript{7,8,g}

\textsuperscript{1} Department of Theoretical Physics, Tata Institute of Fundamental Research, Mumbai 400005, India
\textsuperscript{2} Bethe Center for Theoretical Physics, Physikalisches Institut der Universität Bonn, Nussallee 12, 53115 Bonn, Germany
\textsuperscript{3} National Institute for Theoretical Physics and School of Physics, University of the Witwatersrand, Johannesburg, South Africa
\textsuperscript{4} Instituto de Física Corpuscular, IFIC-UV/CSIC, Valencia, Spain
\textsuperscript{5} Faculty of Physics, University of Warsaw, Warsaw, Poland
\textsuperscript{6} Institute for High Energy Physics, Austrian Academy of Sciences, Vienna, Austria
\textsuperscript{7} Department of Physics, National Tsing Hua University, Hsinchu 300, Taiwan
\textsuperscript{8} Asia Pacific Center for Theoretical Physics (APCTP)-Headquarters San 31, Hyoja-dong, Nam-gu, Pohang 790-784, South Korea

Received: 13 July 2021 / Accepted: 7 October 2021 / Published online: 3 November 2021
© The Author(s) 2021

Abstract Long-lived particles have become a new frontier in the exploration of physics beyond the Standard Model. In this paper, we present the implementation of four types of long-lived particle searches, viz. displaced leptons, disappearing track, displaced vertex with either muons or with missing transverse energy, and heavy charged tracks. These four categories cover the signatures of a large range of physics models. We illustrate their potential for exclusion and discuss their mutual overlaps in mass-lifetime space for two simple phenomenological models involving either a $U(1)$-charged or a coloured scalar.

1 Introduction
The Large Hadron Collider (LHC) started operations over a decade ago and a large number of searches for physics beyond the Standard Model (BSM) have been performed. In particular, the general-purpose experiments ATLAS and CMS have published over a thousand papers each, minutely testing the predictions of the Standard Model and of more exotic theories. For the simplest cases of new physics consisting of new strongly charged particles, corresponding searches already place limits on the BSM masses at about 2–3 TeV \cite{1}, closing in on the maximum reach achievable at a 13 TeV proton collider. We turn our attention therefore to the equally well-motivated but experimentally more challenging cases of new physics where BSM particles are long-lived. A long lifetime may be the natural consequence of a compressed phase space (e.g. in particular dark matter models \cite{2–9}), a suppressed connection to light SM decay products caused by heavy mediators \cite{10}, or direct but feeble couplings to SM particles (possibly resulting from an approximate symmetry) \cite{11–15}. These new kinds of scenarios have raised considerable interest in the community \cite{16} and making the results of associated searches available to reinterpretation in terms
of other theory models has been a frequent request to the CheckMATE collaboration.

With the end of Run 2 of the LHC, both ATLAS and CMS have turned their attention to searches for exotic physics based on new particles with long lifetimes. The signatures of a long-lived particle (LLP) depend on its charges as well as its decay modes and actual lifetime, and they can be fairly complicated to systematically characterize. Several possibilities for LLP signatures indeed exist, based on the LLP decay modes:

- neutral LLP → invisible or neutral stable particles ⇒ missing momentum;
- neutral LLP → charged leptons ⇒ leptons with large impact parameter (i.e. “displaced” leptons);
- neutral LLP → coloured particles ⇒ displaced vertices, or emerging jets;
- stable, charged LLP ⇒ charged track (with its time of flight dependent on mass and boost);
- charged LLP → invisible ⇒ “disappearing” track;
- charged LLP → other charged stable object(s) ⇒ kink-track or displaced vertex.

There is a built-in complementarity in different searches simply because particle decay follows an exponential distribution. For example, charged LLPs with intermediate lifetimes will be visible in both disappearing track and heavy charged track searches. Similarly, a neutral particle decaying into quarks mostly in the electromagnetic or hadronic calorimeter will also likely appear as a smaller, simultaneous signal in the displaced vertex (decay in the tracker) and emerging jet (decay in the hadronic calorimeter) searches. Furthermore, the lifetime of a particle in the lab-frame also depends on its boost, which means that the production mechanism can also significantly alter where in the detector the particle decays. The same particle may accordingly result in different signal distributions depending on whether it is produced “directly” or in the decays of a much heavier particle (i.e. with a higher boost). Finally, several decay modes may be open to a single new particle resulting in sensitivity in multiple searches. The identification of the underlying physics therefore requires a full coverage in terms of the lifetime of new particles.

As we can see, the identification of an LLP is highly complicated and so far, there are no standard algorithms like those available for the identification of standard objects, such as leptons, $b$- or $t$-tagged jets, etc. Consequently, it is not always clear how the results of a dedicated LLP search can be “reinterpreted” for a physics model that differs from the tested one, though it displays a priori similar signatures. A detailed study of models capable of LLP signatures, the reinterpretation struggles and recommendations have been detailed in the community study [16]. In this present work, we use the signal efficiencies published by the experiments in order to implement five searches in the CheckMATE reinterpretation package. The current searches use the existing respective detector implementations for ATLAS and CMS experiments in Delphes. It should also be possible to implement dedicated searches from experiments like FASER [17], CODEX-b [18] or even proposed experiments like MATHUSLA [19] if a corresponding DELPHES module or efficiency parametrisations become available.

CheckMATE [20, 21] is a public tool that allows the reinterpretation of a wide variety of ATLAS and CMS results for new physics models in a coherent and cohesive manner. It consists of an engine written in C++ that runs each analysis cut-by-cut in order to assess the final number of expected events satisfying the requirements of the corresponding analysis. The engine is also capable of using external libraries like MADGRAPH [22] and PYTHIA 8 [23] in order to generate events, while the detector simulation is performed by DELPHES [24]. The User Interface and the statistical analyses are provided by a collection of Python scripts (including the AnalysisManager [25] that guides the users through the implementation of their own analyses).

In Sect. 2, we briefly summarize the main ingredients of the LLP recast, referring the reader to the appendix for a more complete description. Then, in Sect. 3, we illustrate the performance of the implemented searches in two simple models with LLPs, and discuss their complementarity. Conclusions and a brief outlook are proposed in Sect. 4. The appendix consists of a short guide for the user, as well as a more detailed presentation of the implemented LLP searches.

2 Implementation of long-lived particle searches

Below we offer a brief description of the implemented LLP searches and a comparison with experimentally published results. This includes the 8 and 13 TeV versions of the CMS displaced lepton search [26, 27], two different displaced vertex searches [28, 29], the 13 TeV ATLAS disappearing track [30] and heavy charged particle track [31] searches. Together, these searches are capable of probing a wide range of parameter space. Details of the implementation are available in the appendix. Technically, each analysis is encapsulated in a detector-specific “analysis handler” class which provides special functions and efficiencies specific to the detector in question. We deliberately separate the analysis handlers for long-lived particle searches from those used in prompt searches: this accounts for the fact that the implemented prompt searches do not in fact use any decay length infor-

---

1 Internal running of Monte Carlo processes or use of MADGRAPH interface in CheckMATE is currently compatible only with the 8.2 series of PYTHIA 8.
mation and all particles denoted stable by the Monte Carlo generator are then clustered into jets based on their kinematics only.\footnote{For this reason, we advise user discretion when applying prompt search limits to models with LLPs. It may be possible to find a conservative limit from prompt searches by e.g. removing decay products of LLPs from the Monte Carlo events beforehand. However each case needs to be evaluated separately and we do not provide a built-in solution for that reason.}

2.1 Displaced lepton searches

The displaced lepton searches \cite{26,27} look for two high-$p_T$, isolated leptons ($\ell$) with large impact parameter relative to the primary vertex. The benchmark used for this search is motivated by R-parity violating (RPV) supersymmetry \cite{13,32} where a top-squark ($\tilde{t}_1$) decays via the lepton-number-violating LQD operator as $\tilde{t}_1 \to \ell b$. The leptons thus produced have large $p_T$ and are well isolated. The two searches implemented here correspond to 8 TeV \cite{26} and 13 TeV \cite{27} versions of the CMS displaced supersymmetry search. The identification and fiducial acceptances are provided on generator-level events. We therefore reproduce the Monte-Carlo production process for validation of the search. Corresponding details are provided in Appendix B.

The event selection for both 8 TeV and 13 TeV analyses was performed in two stages. The first stage (i.e. preselection) selects events with exactly one electron and one muon with opposite electric charges, each expected from the decay of a different top squark. Further selection cuts and isolation requirements are then applied. In the second stage, the events are classified into three signal regions (SR) corresponding to increasing ranges of the leptonic impact parameter $d_0$.

The validation results are shown in Fig. 1, with the experimental exclusion limits displayed in blue, while the recast produces the black exclusion bound. A reasonable agreement is observed at 8 TeV in Fig. 1a. The situation at 13 TeV is somewhat more subtle.

Indeed, as efficiencies have been provided by the experimental collaboration for the 8 TeV, but not for the 13 TeV search – in particular, the considered ranges of $p_T$ and $d_0$ do not match and the modelling of efficiencies in this latter case appears as an important assumption in the recast. An attempt to validate the 13 TeV search using the 8 TeV efficiencies results in poor agreement with the numbers in the signal region at large impact parameters, and therefore in a much stronger expected limit for high values of the LLP lifetime $\tau_\ell$ (see the dashed black curve in Fig. 1b). A simple linear interpolation performs rather poorly as well. We therefore make a conservative estimate of this detector effect by adding a single bin for the $d_0$ range (20 mm - 100 mm) and determine the associated efficiency via a $\chi^2$-fit of the expected number of events in all three signal regions. The outcome of this procedure is displayed in Fig. 1b as a solid black line and shows a considerably improved agreement with the expected exclusion limits. Exact numbers in each of the signal regions are produced in Appendix B.

2.2 Displaced vertex searches – DV + MET

This ATLAS search \cite{28} looks for high-mass displaced vertices (DV$s$), reconstructed from five or more tracks. Large missing transverse momentum is also required. The outcome of 32.8 fb$^{-1}$ of 13 TeV collision data is a yield consistent with the expected background.

The template considered by the ATLAS collaboration consists in the (strong) production of a pair of long-lived gluinos ($\tilde{g}$), then decaying into light quarks and stable neutralinos. Heavy squark mediators result in suppressed gluino decay widths. Recast instructions were provided in \cite{33} and include a preselection at generator-level, followed by the application of parametrized efficiencies. Previously to our implementation, this strategy has been applied with success by the publicly available codes \cite{34,35} (see also Contribution 22 in \cite{36}). Details on the implementation are provided in Appendix C.

For the validation, we considered the limits on the gluino production cross-sections presented in the ATLAS paper. In a first scenario, the long-lived gluino and the neutralino LLP are separated by a wide mass-gap, with the neutralino fixed at 100 GeV while the gluino takes mass of $m_{\tilde{g}} = 1.4$ TeV or 2 TeV. The LLP lifetime is varied between $\tau_\ell = 0.003$ ns and 50 ns. The results from the recast search are shown in Fig. 2, using two statistical approaches: the simplified evaluation of CheckMATE defining a ratio $r$ (blue curve) – see Eq. (1) of \cite{21} – and the full $p$ value analysis (red curve); both return very similar limits. The statistical uncertainty in the simulation is below percent level (10$^6$ events are generated at each point), except for the end point $\tau_\ell = 0.003$ ns, where it reaches 3–4%. The 95% CL limits from the experimental analysis is shown in black. We observe a general qualitative agreement. In the lower row of plots of Fig. 2, the experimental observed limits are normalized to the limiting cross-sections of the recast procedure (with $r$-approach in blue and $p$-values in red). Quantitatively, we find that the bounds agree within 20% of the cross-section value, with outliers at up to 50% discrepancy for small lifetimes.

Nevertheless, this apparent success of the recast strategy with efficiencies applied on truth-level objects needs to be tempered as it seems to perform worse in the case of a compressed spectrum. This was confirmed to us by the authors of \cite{34,35}. A more detailed comparison is provided in the appendix.
Fig. 1 A Comparison of the exclusion limits on the Displaced Lepton search provided by CMS with those obtained from CheckMATE (left: 8 TeV, 19.7 fb$^{-1}$; right: 13 TeV, 2.6 fb$^{-1}$).

Fig. 2 Validation of the DV+MET search in the scenario with large mass-splitting for two different benchmarks (left $m_{\tilde{g}}=1.4$ TeV, right: $m_{\tilde{g}}=2$ TeV). The bottom panel in both cases shows the ratio of the published ATLAS exclusion to the one obtained by CheckMATE. There is no observable difference between using the full CLs method and an exclusion based on the ratio ($r$) of cross section from CheckMATE of events passing all cuts to the 95% upper limits on cross sections published by ATLAS.
2.3 Displaced vertex searches – DV + μ

In this section we discuss a search for massive, long-lived particles decaying to final states with a DV and an energetic muon [29]. The search analyzed 139 fb\(^{-1}\) of data collected by ATLAS at the centre of mass energy 13 TeV.

The benchmark process considered by the experiment was pair production of top squarks followed by the RPV decay into a light quark and a muon. Other physics scenarios, for example, long-lived lepto-quarks, right-handed neutrinos or long-lived electroweakinos in RPV, could result in similar signals including a DV and a muon. In Sect. 3 we apply this search to sbottom pair-production followed by the RPV decay.

The event selection defines two mutually exclusive trigger-based signal regions: \(E_{\text{T}}^{\text{miss}}\)-Trigger SR and Muon Trigger SR. The former requires significant missing transverse momentum (> 180 GeV), while the latter is recorded with the muon trigger and has low (< 180 GeV) transverse momentum. Additionally, at least one displaced vertex is required to be present in the fiducial region. There is no explicit requirement for a signal muon to originate from the reconstructed vertex.

The search was validated using a benchmark RPV-supersymmetric (SUSY) model for the process \(pp \rightarrow \tilde{\chi}_{1}^{\pm} \tilde{\chi}_{1}^{0}\). In Fig. 3 we show a comparison of the ATLAS result and CheckMATE recasting in the stop lifetime-mass plane, \(\tau_{\tilde{t}}-m_{\tilde{t}}\). The yellow band shows a 2-sigma range of the ATLAS expected exclusion limit, the blue solid line is the ATLAS observed exclusion while the blue dashed the ATLAS expected exclusion limit, the black solid line shows an recasting in the stop lifetime-mass plane, \(\tau_{\tilde{t}}-m_{\tilde{t}}\). The yellow band shows a 2-sigma range of the ATLAS expected exclusion limit, the blue solid line is the ATLAS observed exclusion while the blue dashed the ATLAS expected exclusion limit, and the black solid line shows an exclusion line obtained with CheckMATE. Generally a good agreement is visible between the ATLAS (red) and the CheckMATE-derived (dashed blue) limits for both chargino and stau scenarios. More details can be found in Appendix E.

2.4 Heavy charged particles searches

In this section we focus on the search for heavy charged long-lived particles performed by the ATLAS experiment using a data sample of 36.1 fb\(^{-1}\) of collisions at 13 TeV [31]. In our implementation we cover searches for long-lived charginos and sleptons.

The ATLAS collaboration reported no significant excess of observed data events above the expected background in this search. Thus, the collaboration have published upper limits at 95% confidence level on the cross-sections for stau and chargino production for specific benchmark models. These limits have been obtained applying the CLs prescription [37].

For the validation of our implementation in CheckMATE we have employed HistFitter [38] to estimate the CLs while running 10\(^5\) toy experiments, given the low backgrounds for the considered signal regions and assuming a 10% signal uncertainty.

The validation has been performed through the comparison with the observed upper cross-section limits reported by the ATLAS collaboration, as is depicted in Fig. 4. A very good qualitative agreement is visible between the ATLAS and the CheckMATE-derived (dashed blue) limits for both chargino and stau scenarios. More details can be found in Appendix E.

2.5 Disappearing track searches

The ATLAS collaboration presented a search\(^3\) [30] for direct electroweak (EW) gaugino or gluino pair production with wino-like electroweakinos (hence near-degenerate charged and neutral SU(2)-triplet fermions). The chargino decays via \(\tilde{\chi}^{+} \rightarrow \pi^{\pm}\tilde{\chi}^{0}\); the neutralino is stable. The experimental collaboration analysed data based on the integrated luminosity of 36.1 fb\(^{-1}\) recorded between 2015 and 2016. Due to the small mass difference between the two states (which is of the order of the pion mass), the chargino is long-lived and the decay products are entirely invisible to the detector. Thus the chargino appears as a “disappearing” track, i.e. a track that does not reach the outer edges of the tracker detector but stops before. In order to define a trigger isolating the signal from large SM backgrounds, the search additionally demands a large momentum jet from initial-state radiation or four jets originating from the gluino decay. The observed number of events is consistent with the SM expectations and constraints for wino-like charginos with a lifetime of 0.2 ns, and mass up to 460 GeV are derived. In the strong produc-

\(^{3}\) HEPDATA accessed from https://doi.org/10.17182/hepdata.78375.v5.
Fig. 4 Validation of the HCP search – observed upper cross-section limits using the best expected signal region for chargino pair production (left) and stau pair production (right). ATLAS results are shown as a solid red line while CheckMATE predictions are in form of a dashed blue line. The theory prediction is displayed as a dot-dashed black line.

**Table 1** Summary of the selection criteria for signal events for direct electroweakino production and the strong channel channel where the chargino is produced in gluino decays

|                 | EW SR                        | Strong SR                   |
|-----------------|------------------------------|-----------------------------|
| At least one jet with $p_T > 140$ GeV | $p_T(j_1) > 100$ GeV, $p_T(j_2) > 50$ GeV, $p_T(j_3) > 50$ GeV |
| $E_T^{\text{miss}} > 140$ GeV          | $E_T^{\text{miss}} > 150$ GeV |
| $\Delta \Phi (E_T^{\text{miss}}, \text{jets}(p_T > 50 \text{ GeV})) > 1.0$ | $\Delta \Phi (E_T^{\text{miss}}, \text{jets}(p_T > 50 \text{ GeV})) > 0.4$ |
| Cuts on charged LLP with $p_T > 100$ GeV | Cuts on charged LLP with $p_T > 100$ GeV |

**Table 2** Cutflow comparison for a chargino produced in direct electroweak production with $(m_{\tilde{\chi}^\pm}, \tau_{\tilde{\chi}^\pm}) = (400 \text{ GeV}, 0.2 \text{ ns})$

|                     | CM $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$ | CM $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ | CM $\tilde{\chi}_1^- \tilde{\chi}_1^0$ | CM all channels | ATLAS |
|---------------------|---------------------------------------------|-----------------------------------------|-----------------------------------------|-----------------|-------|
| Trigger             | 445.1                                       | 624.0                                   | 274.4                                   | 1343.5          | 1276  |
| Lepton Veto         | 423.4                                       | 608.5                                   | 267.3                                   | 1308.2          | 1181  |
| MET and jet requirements | 164.2                                     | 229.6                                   | 101.0                                   | 494.8           | 579   |
| EW SR               | 5.2                                         | 4.4                                     | 1.6                                     | 11.2            | 13.5  |

We use the ATLAS benchmark points as the test case scenarios which correspond to $(m_{\tilde{\chi}_1^\pm}, \tau_{\tilde{\chi}_1^\pm}) = (400 \text{ GeV}, 0.2 \text{ ns})$ for the EW signal region and $(m_{\tilde{\chi}_1}, m_{\tilde{\chi}_1^\pm}, \tau_{\tilde{\chi}_1^\pm}) = (1600 \text{ GeV}, 500 \text{ GeV}, 0.2 \text{ ns})$ in the strong signal region. The results for the EW and strong SR are summarized in Tables 2 and 3, respectively and our recast results show satisfactory agreement with the public ATLAS results.

We did not validate the disappearing track search in a grid scan with ATLAS exclusions, since the event generation requires matched events with two additional partons in the final state in order to reproduce the ATLAS cutflows. As only a few events would pass all selection cuts, such a scan would be costly to perform and we thus opted against it.

### 3 Performance and interplay in LLP scenarios

In this section, we consider two simple LLP scenarios that put forward the complementarity of the implemented searches.
Table 3 Cutoflow comparison for a chargino produced in strong production channel with \((m_{\tilde{b}_1}, m_{\tilde{b}_1}, \tau_{\tilde{b}_1}) = (1600 \text{ GeV}, 500 \text{ GeV}, 0.2 \text{ ns})\) in the high–MET region.

|                 | CM    | ATLAS |
|-----------------|-------|-------|
| Trigger         | 289   | 285   |
| Lepton Veto     | 277   | 278   |
| MET and jet requirements | 216   | 202   |
| Strong SR       | 11    | 11    |

3.1 Electroweak LLP

We first consider a model addressing the case of electroweakly produced LLPs. It extends the SM with a scalar (\(\phi\), charged under \(U(1)_Y\) and sharing the gauge quantum numbers of a right-handed lepton, and a SM-singlet Dirac fermion (\(\chi\)). An extra \(Z_2\)-symmetry under which both \(\phi\) and \(\chi\) are charged further constrains their interactions. The scalar is produced in pairs via the Drell–Yan process \(pp \to \phi^* \phi\). After that it decays according to \(\phi \to \ell \chi\), mediated by the Yukawa coupling \(y_{\ell} \phi \bar{\ell} \chi + \text{h.c.}\). The singlet fermion is assumed to be stable. The model has been implemented in \textsc{Pythia} 8. The Lagrangian is simply

\[
\mathcal{L} = \frac{1}{2} D_{\mu} \phi D^{\mu} \phi - \frac{1}{2} m_{\phi}^2 \phi^2 + \bar{\chi} (i \gamma^\mu \partial_\mu) \chi - m_{\chi} \chi - \sum_{\ell} (y_{\ell} \phi \bar{\ell} R \chi + \text{h.c.}) \]  

For small values of the Yukawa coupling, the scalar \(\phi\) is long-lived and could be visible in the charged track searches. In the case where the lifetime is too short to leave the required hits in the tracker modules, the displaced lepton search might detect the products of the decay. To examine the complementarity of these searches, we set \(y_e = y_\mu\), ensuring that the targets of the displaced lepton search are actually produced. As the lifetime is determined by the smallness of the Yukawa coupling and is fairly independent of the mass of \(\chi\), we set \(m_\chi = 10 \text{ GeV}\). These assumptions of course mean a focus on a very specific scenario.

The exclusion results obtained with \textsc{CheckMATE} in the lifetime-mass plane are shown in Fig. 5. The CMS displaced-lepton searches appear sensitive to the considered scenario for LLP lifetimes up to \(\sim 1 \text{ ns}\) and masses in the range 100–500 GeV. The Heavy Charged Particle search from ATLAS impacts the high–MET region \(c\tau > 1 \text{ m}\), in a comparable range of masses. The disappearing track search does not show any sensitivity, due mainly to cross-sections much weaker in the considered model than the targeted range of the search: the scalar production cross section is indeed suppressed compared to the fermionic one. Moreover, the original disappearing track final state consists of multiple final states (all wino final state configurations), e.g. the NLO cross section for scalar leptons of mass 125 GeV is about 0.045 pb whereas the combined wino pair production cross section is about 1 pb. Finally, the cross section is further suppressed since the disappearing track search basically triggers on monojet events which requires a hard recoil of the scalar pair against a hard jet. As a consequence, the current model cannot be probed with the disappearing track search which is focused on mass degenerate wino-like electroweakinos. Still, it is obvious that two LLP searches aiming at quite different signatures can interplay and lead to complementary exclusion bounds. Current prompt limits on pair production of scalars that decay to electron or muon with missing energy are at 250 GeV with full Run 2 data [40].

3.2 Strongly-interacting LLP

Just as EW-charged LLPs manifest in the form of track-based signatures, strongly charged LLPs result in jets originating from a secondary vertex. A possible example is provided by the minimal supersymmetric standard model, where small R-parity violating couplings [32] open up decay channels of coloured R-odd particles. We shall focus in particular on the LQD coupling \(\lambda'_{223}\) where the subscripts refer to an interaction between the second lepton doublet superfield, the second left-handed quark doublet superfield and the right-handed bottom superfield. This results in various possible decay channels like \(\tilde{b}_1 \to \mu c\) or \(\tilde{b}_1 \to \nu s\), where \(\tilde{b}_1\) is the lightest bottom squark, which we assume to (nearly) coincide with the long-lived right-handed sbottom.

To test this scenario, we employ a simplified supersymmetric model where all the new-physics particles take a mass at the 10 TeV scale, beyond the discovery reach of the LHC, with the exception of the lightest bottom squark of right-handed type \(\tilde{b}_1\), whose mass is scanned over in the electroweak-TeV range. The lifetime of \(\tilde{b}_1\) is then entirely determined by the size of the coupling \(\lambda'_{223}\) and can be varied freely. We generate samples of \(10^6\) events for \(pp \to \tilde{b}_1 \tilde{b}_1^*\) using \textsc{Pythia} 8. The production cross section is normalized to the NNLOapprox + NNLL values provided by the LHC cross-section Working Group [41].

Several LLP searches are potentially sensitive to this scenario. The decay channel \(\mu + \text{jet}\) is an obvious target for the DV+muon analysis, while the decay into neutrino + jet enters the scope of the DV+MET search. Finally, the bottom squark could be detected as a heavy long-lived charged particle. However, we do not compute limits from the corresponding search due to known uncertainties in the hadronization of such long-lived strongly charged scalars. In addition, the disappearing track search is insensitive here because the displaced jets produced from the sbottom decays are hard, as a general rule.
Upper limits in the lifetime-mass plane (left) and coupling-mass plane (right) of the LLP scalar $\phi$. We do not get limits from the disappearing track search due to a veto on leptons in the event.

Limits in the mass-lifetime (left) and mass-$\lambda(=\lambda'_{223})$ plane of the LLP sbottoms.

The limits obtained with CheckMATE are presented in Fig. 6 in the plane corresponding to the lifetime and mass of the bottom squark. We observe that, in this configuration where muon and MET productions are set equal, the DV+muon search proves slightly more competitive than the DV+MET analysis (which is consistent with the respective limits placed on the cross-sections by these searches in their respective benchmark scenario). The most constraining limits are placed on lifetimes $\sim 1$ ns and exclude sbottom masses up to $\sim 1.6$ TeV. This is more competitive than limits from the R-parity conserving scenario with $b \rightarrow b \tilde{\chi}^0$, which places a limit of 1270 GeV for a massless $\tilde{\chi}^0$ [42].

Two types of prompt searches would also constrain the same model in a different parameter space searches for promptly-decaying leptoquarks, or usual sbottom limits (in the presence of some light LSP with the decay $b \rightarrow c \tilde{\chi}^+ \rightarrow c \ell \nu \tilde{\chi}^0$). There have not been direct searches for sbottoms using this topology. Limits on squarks in the $2\ell+2$jets+MET or in 2jets + MET are both expected to be much smaller than the standard topology ($b \rightarrow b \tilde{\chi}^0$). Limits on standard sbottom decay with full run 2 data are currently between 600 and 1270 GeV [42] depending on the mass of the final invisible particle. Although there have been searches for third generation leptoquarks, they focus on decays into tops [43,44] (requiring e.g. b-tagged jets) and therefore do not apply directly to our model.

4 Comments and outlook

In this paper, we present the implementation of a new class of analyses in the CheckMATE package, dedicated to long-lived particle searches, which can be used to reinterpret experimental limits on new physics models. We demonstrate the interplay of these searches in detecting strongly or weakly charged LLPs and obtain limits comparable to prompt limits in certain ranges of lifetime. In addition, these searches provide a way to probe couplings that can even be much smaller than those currently observable via low-energy intensity-frontier experiments, e.g. by measuring meson decays.

For our two scalar models, we find that the electroweak model can be constrained up to a mass of 480 GeV with the CMS dilepton search for lifetimes under 10 cm. The charged track search is able to set bounds up to about 400 GeV for large lifetimes, greater than 10 m.

We find an important gap in the search for electroweakly charged LLPs decaying to leptons in the intermediate lifetime range. Since the disappearing track search employs a lepton
veto, it is not possible currently to understand whether models with decays into leptons can be observed (as the leptons emerge with potentially very large impact parameters and not from the primary vertex). It would be fruitful to both improve the experimental search criteria as well as to provide efficiency for lepton identification based on $p_T$ and $d_0$ to remedy this situation.

For the strongly charged LLP, our limits are stronger than typical SUSY searches for particles with the same quantum numbers because of the much smaller backgrounds in these exotic searches. An important gap here is that interaction of these particles with the detector material results in the efficiencies being dependent on some more unknown parameters than just the mass and lifetime. This is highlighted in Appendix C. Also, this same issue prevents us from naively applying the charged track search to strongly interacting LLPs where charge exchange with the detector material becomes important.

Due to the absence of standardized detector objects in such searches, each re-interpretation heavily depends on the parametrized efficiencies published by the experiments. The validation of the five searches considered here has shown that this method gives relatively good results, which should allow the recast of the experimental limits for a very diverse range of models sharing an LLP as a common feature. However, since the efficiencies do rely on identifying the right truth-level particle via a user input and PDG code (possibly assigned ad hoc by the user for new particles), we do advise user vigilance when using these results.

Users can implement their own versions of an LLP analysis using the AnalysisManager, which simplifies the setting up of detector parameters, stores the expected and observed events reported by the experiment in the correct format for future use in statistical calculations and provides a skeleton C++ analysis code with access to all detector objects. Currently, searches based on ionisation, like those for monopoles or multi-charged objects, are not possible to implement because the relevant efficiencies are not publicly available. However, if the efficiency tables based on mass, charge and momentum were to be provided by the experiments in the future, it would be possible to implement them as well, without further difficulties.

Acknowledgements We thank Giovanna Cottin and Andre Lessa for useful discussions. R. RdA acknowledges partial funding/support from the Elusives ITN (Marie Skłodowska-Curie grant agreement No 674896), the “SOM Sabor y origen de la Materia” (FPA 2017-85985-P). Z. S. W. is supported partly by the Ministry of Science and Technology (MoST) of Taiwan with grant number MoST-109-2811-M-007-509, and partly by the Ministry of Science, ICT & Future Planning of Korea, the Pohang City Government, and the Gyeongsangbuk-do Provincial Government through the Young Scientist Training Asia Pacific Economic Cooperation program of the Asia Pacific Center for Theoretical Physics. Z. S. W. also thanks the Witwatersrand University for hospitality where part of this work was initiated. KR is supported by the National Science Centre, Poland under Grants: 2016/23/G/ST2/04301, 2015/18/M/ST2/00518, 2018/31/B/ST2/02283, 2019/35/B/ST2/02008, and the Norwegian Financial Mechanism for years 2014-2021. Grant DEC-2019/34/H/ST2/00707. F.D. acknowledges support of the BMBF Verbundprojekt 05H2018 and the DFG Grant SFB CRC-110 Symmetries and the Emergence of Structure in QCD.

Data Availability Statement This manuscript has associated data in a data repository. [Authors’ comment: The latest version of CheckMATE can always be downloaded from https://github.com/CheckMATE. Documentation and validation notes can be found at https://checkmate.hepforge.org/.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP³.

Appendix A: Usage

In this section, we demonstrate the usage of CheckMATE with LLP searches on an example, describing the initialization and running of a test. We assume that CheckMATE has already been successfully installed.⁴ We then focus on a simplified SUSY scenario where only the gluino and the lightest neutralino are kinematically accessible, while the rest of the R-odd spectrum is heavy and decoupled: we employ the same benchmark point as in Sect. 2.2, with $m_{\tilde{\chi}^1}=0.1$ TeV, $m_{\tilde{g}}=1.4$ TeV, and $\tau_{\tilde{g}}=0.1$ ns. The gluinos are produced at the LHC with 13 TeV center-of-mass energy through strong interactions: the corresponding events are generated by calling PYTHIA 8 internally. At this point, the CheckMATE user needs to provide input to the program, either via a parameter card or via command line. We consider the minimal input through text file (parameter card) below. For the command line call, we refer the reader to Ref. [25].

⁴ For more detailed information on CheckMATE operations, please refer to Refs. [20, 21].
The text command file is structured in blocks, which are introduced by expressions between brackets and contain one or more Key: Value pairs. 

[Parameters] provides general settings that are common to all processes in the CheckMATE run. In the example above, the run is called ExampleRun, which also determines the name of the result directory. The Analyses parameter selects the list of analyses against which the generated events are tested: a list of predefined identifiers involving LLP searches is provided in Table 4. Only the displaced-vertex+MET search is kept in the file above. The PDG code of the LLP (gluino) is defined in the longlivedPIDs entry, while that of exotic stable electroweakly-interacting particles (the neutralino LSP in our case) is fed via invisiblePIDs. In the current CheckMATE version, only a single longlivedPIDs and invisiblePIDs are supported. Finally, the path to the SLHA [45,46] file defining the spectrum is provided after the entry SLHAFile.

The following [X] blocks list individual production processes (only one in our example), with X the corresponding (and freely chosen) identifier. Pythia8Process defines the considered production mode. MaxEvents sets the number of generated events. XSect contains the corresponding cross-section value at 13 TeV center-of-mass energy, taken from the LHC cross-section Working Group at NNLOapprox+NNLL [47].

CheckMATE is executed with the following command:

```
$CMDIR/bin/./CheckMATE checkmate_example.in
```

The program responds with a summary of the inputted settings for the considered run and prompts confirmation from the user. The following settings are used:

Analyses: atlas_1710_04901

Processes:

- Process Name: 1710_04901_example_run
- Input Cross section: 0.0284 PB
- Associated event files and/or Monte-Carlo generation runs:
  - Pythia8 Events
    - internal identifier: 1710_04901_example_run
    - simplified SUSY process: p p > go go
    - at most 10000 events are generated and analysed

Output Directory:

$CMDIR/results/ExampleRun

Additional Settings:

- SLHA file path/1710_04901_test_gluino.slha will be used for event generation
- The following PIDs will be considered as invisible for the detector: 1000022
- The following PIDs will be considered as long lived for the detector: 1000021

Is this correct? (y/n)

In this simple example, CheckMATE tests against the displaced-vertex+MET search from ATLAS. Returning y launches the (Pythia+(+CheckMATE run. During its course, updates of the test status are printed on the screen. Finally, the program displays the result summary:

```
Evaluating Results
Test: Calculation of approximate (fast) likelihood given in results folder
Result: Excluded
Result for r: 150.498232611
Analysis: atlas_1710_04901
SR: SR1
```

In agreement with the ATLAS result presented in Ref. [28], or the left plot in Fig. 2, the benchmark point is thus excluded ($r \gg 1$) by the only analysis (atlas_1710_04901) considered in the CheckMATE test, in the signal region SR1.
Table 4 Table of implemented LLP analyses and their identifiers that can be used in the CheckMATE run card

| Analysis                                      | CheckMATE identifier                       |
|-----------------------------------------------|--------------------------------------------|
| ATLAS disappearing track EW                  | atlas_1710_04901_ew                        |
| ATLAS disappearing track QCD                 | atlas_1710_04901_qcd                       |
| ATLAS heavy charged track (EW only)           | atlas_1902_01636                           |
| ATLAS displaced vertex (DV with lepton veto) | atlas_1710_04901                           |
| ATLAS muon plus displaced vertex             | atlas_2003_11956                           |
| CMS 8 TeV displaced vertex                   | cms_1409_4789                              |
| CMS 13 TeV displaced vertex                  | cms_pas_exo_16_022                         |

Note concerning the compatibility of LLP and prompt searches in CheckMATE

The explicit partitioning between prompt and LLP searches is in particular motivated by the absence of in-built veto against LLP phenomena in the implemented prompt searches, leading to inconsistent results when testing events with particles having displaced production or decay points. Obviously it would be possible to conservatively test LLP spectra against prompt searches by calculating the fraction of LLPs decaying before hitting the tracker in the simulated events, then multiplying the $r$-value obtained with CheckMATE by this fraction. Such an approach was employed in a recasting tool SMODELS [48]. Yet, such a test against prompt searches would still require a separate CheckMATE run as compared to the test against LLP searches.

In fact, the fundamental reason for this separation is that the standard efficiencies on reconstructed detector level objects (such as jets, leptons, missing transverse energy) used in prompt searches do not necessarily apply to the case of LLP searches. Individual LLP analyses e.g. employ truth-level information (i.e. properties of the MC event generation prior to the fast detector simulation), with recasting efficiencies provided by the experimental collaborations explicitly applying on truth-level objects. We provide a new directory in the top-level data directory called tables where any efficiency tables can be stored. ROOT file format is preferred as it is easily available from experimentalists but any format that can be read directly from the analysis code in is allowed.

There are also situations where two different LLP searches should not be run simultaneously. For example the disappearing track search has two different efficiency tables published by the ATLAS disappearing track search for tracks due to electro-weak particles and strong particles. Only one of these should be on at a time. In general, we suggest that users build up by hand the list of LLP analyses that they wish to run simultaneously. To facilitate this, Table 4 lists the CheckMATE identifiers of all the implemented LLP analyses.

Appendix B: Implementation details: displaced lepton

We use simulated Monte Carlo samples to evaluate the acceptance of the search regions. The samples are generated for the process $pp \rightarrow t_1t_1^*$, for a top squark mass of 500 (700) GeV, using MADGRAPH5 [22] to produce the LHE file for both the 8 TeV and 13 TeV runs. The top squark further decays via an RPV vertex into electrons and muons with a branching ratio of 0.5 each. Total production cross section is normalized to the stop-pair production cross section at NLO.\footnote{The LHC SUSY cross section working group publishes updated calculations of total cross sections at https://twiki.cern.ch/twiki/bin/view/LHCPHysics/SUSYCROSSSections.} To recover the assumption of lepton universality, we multiply the overall cross section with a factor of 2/3. Events are generated using PYTHIA 8, followed by DELPHES for simulation of the CMS detector acceptance and efficiencies respectively. FASTJET 3 [49] was used to construct jets, using the anti-kt algorithm [50] with a distance parameter of 0.4. Ten thousand events are generated for the described process, for each of the following proper lifetimes of the top squark: 1 mm, 10 mm, 100 mm. Additionally, for the 13 TeV analysis, a sample is generated for the stop proper lifetime 1 m.

Event selection

Throughout this section, the 13 TeV criteria (wherever they differ) are shown within a bracket following the 8 TeV requirements. At the preselection stage, events with exactly one electron and one muon with opposite charge are singled out. Further conditions on the leptons to pass the detector trigger requirements are as follows:

- $p_T^e > 25(40)$ GeV
- $p_T^\mu > 25(42)$ GeV
- $|\eta| < 2.5(2.4)$

Additional isolation requirements are placed on the leptons by requiring that the sum of $p_T$ of all tracks that fall within some $\Delta R$ of the lepton track are smaller than a fraction $f$ of...
the lepton-pT. Here are the requirements in terms of ($\Delta R$, $f$) for the different cases.

- Muons: (0.5, 0.15)
- Electrons with $1.56(1.57) < |\eta^e| < 2.5(2.4)$: (0.3, 0.065)
- Electrons with $|\eta^e| < 1.44$: (0.3, 0.035)

Electron candidates are rejected if they have $1.44 < |\eta| < 1.56$ due to reduced reconstruction performance of electrons in the overlap region between the barrel and endcap detectors. For the 8 TeV search, there must be no jets within a cone of $\Delta R = 0.5$ around either lepton and the electron and muon must be separated by $\Delta R > 0.5$. There is no jet-isolation requirement for the 13 TeV search.

Events passing the preselection criteria are further classified according to the impact parameters of the leptons. The impact parameter $|d_0|$ is defined as the distance of closest approach of the helical trajectory of the lepton in the transverse plane to the beam axis.

$$|d_0^L| = \frac{|p_{x}^L - p_{x}^L|}{p_T^L}$$

where $p_{x}^L$ and $p_{x}^L$ are the radial components of the lepton’s momentum while $L_{y}$ and $L_{y}$ are the radial components of the decay vertex (position) of the top squark from which the lepton originates. By requiring the leptons to have a large enough impact parameter, one reduces the likelihood of the leptons of having originated from standard model processes. Both leptons are therefore required to have a minimum impact parameter $d_0^L > 0.1(0.2)$ mm. Three signal regions are defined based on these impact parameters:

- SR1: $d_0 > 0.2$ mm for both leptons and $d_0 < 0.5$ mm for at least one;
- SR2: $d_0 > 0.5$ mm for both leptons and $d_0 < 1$ mm for at least one;
- SR3: $d_0 > 1$ mm for both leptons.

Furthermore, events in which either lepton has an impact parameter $d_0 > 20(100)$ mm is rejected, in order to ensure that the leptons originate within the pixel layer.

**Comparison**

To compare the results of this analysis with the values determined by the CMS collaboration [26,27], we define the percentage difference as

$$\text{diff} = \frac{|n_{\text{CheckMATE}} - n_{\text{CMS}}|}{n_{\text{CMS}}}$$

where $n$ is the number of events that pass all selections in each signal region.

The results as well as the comparison with the CMS analysis are shown in Tables 5 and 6.

The analysis was conducted over top squark masses ranging from 300 (400) GeV to 850 (900) GeV, and top squark lifetimes $c\tau$ from 0.1 mm to 1 m. A simultaneous counting experiment was performed on the three bins of the three signal regions, and the resulting exclusion contours compared to

**Table 5** Comparison of the number of expected events in the three search regions for the displaced lepton search with the CMS detector at 8 TeV with CheckMATE (CM), for the process $pp \rightarrow t_1\bar{t}_1^*$, with $M_{t_1} = 500$ GeV and stop quark lifetimes of 1, 10 and 100 mm. The CMS results are presented here with statistical and systematic uncertainties combined in quadrature.

| $c\tau$ (mm) | Search Region 1 | Search Region 2 | Search Region 3 |
|---------------|----------------|----------------|----------------|
|               | CM   | CMS | Diff | CM   | CMS | Diff | CM   | CMS | Diff |
| 1              | 27.2 | 30.1 ± 3.53 | 9.63% | 6.86 | 6.54 ± 1.21 | 4.89% | 1.26 | 1.34 ± 0.28 | 5.97% |
| 10             | 37.6 | 35.3 ± 6.25 | 6.52% | 32.0 | 30.3 ± 5.35 | 5.61% | 50.0 | 51.3 ± 9.06 | 2.53% |
| 100            | 4.78 | 4.73 ± 0.88 | 1.06% | 6.13 | 5.57 ± 1.03 | 10.0% | 26.5 | 26.3 ± 4.65 | 0.760% |

**Table 6** Comparison of the number of expected events in the three search regions for the displaced lepton search with the CMS detector at 13 TeV with CheckMATE (CM), for the process $pp \rightarrow t_1\bar{t}_1^*$, with $M_{t_1} = 700$ GeV and stop quark lifetimes of 1, 10 and 100 mm. The CMS results are presented here with statistical and systematic uncertainties combined in quadrature.

| $c\tau$ (mm) | Search Region 1 | Search Region 2 | Search Region 3 |
|---------------|----------------|----------------|----------------|
|               | CM   | CMS | Diff | CM   | CMS | Diff | CM   | CMS | Diff |
| 1              | 3.85 | 3.80 ± 0.2 | 1.24% | 0.99 | 0.94 ± 0.06 | 4.8% | 0.19 | 0.16 ± 0.02 | 20% |
| 10             | 5.18 | 5.20 ± 0.4 | 0.38% | 4.6 | 4.1 ± 0.3 | 11% | 7.6 | 7.0 ± 0.3 | 8.0% |
| 100            | 0.86 | 0.80 ± 0.1 | 7.0% | 1.04 | 1.0 ± 0.1 | 3.97% | 5.4 | 5.8 ± 0.2 | 7.5% |
| 1000           | 0.03 | 0.009 ± 0.005 | ~ 260% | 0.05 | 0.03 ± 0.01 | 59% | 0.45 | 0.27 ± 0.03 | 67% |
those provided by CMS \cite{26,27}. Figure 1 of Sect. 2.1 shows these exclusion contours for the 8 TeV and 13 TeV analyses. It is observed that the CheckMATE 8 TeV exclusion contours agree well with CMS, except for $c\tau$ between about 7 mm to 70 mm, where the CheckMATE limits are stronger.

For the 13 TeV analysis, the absence of available efficiencies complicate the recast. The use of the 8 TeV values results in satisfactory agreement at small lifetimes up to about 5 mm. However, as the 13 TeV search extends the allowed $d_0$ up to 100 mm, it is necessary for the recast to assess the efficiency in the $d_0$-bin 20–100 mm. The lepton $d_0$ efficiency is linearly extrapolated at 22 mm using the values in the range [12 mm, 20 mm] for both electrons and muons. Keeping the electron $d_0$ efficiency in the added bin constant, we vary the muon $d_0$ efficiency in the added bin and minimize the $\chi^2$ fit over all resulting signal regions for the benchmark stop lifetimes 100 mm and 1000 mm. The minimum $\chi^2$ value occurs for a muon $d_0$ efficiency of 0.01 in the added bin. Using this value, we similarly minimize the $\chi^2$ fit for electron $d_0$ efficiency values in the added bin. This minimum occurs for an electron $d_0$ efficiency of 0.06. The resulting exclusion plot still exhibits stronger limits than those of the CMS publication, but shows a significant improvement over the exclusion plot obtained with only the 8 TeV lepton reconstruction efficiencies.

Appendix C: Implementation details: displaced vertex and MET

Monte Carlo simulation samples

As a template for this search, the ATLAS collaboration considered the (strong) production of a pair of long-lived gluinos, then decaying into light quarks and a stable neutralino: $pp \rightarrow \tilde{g}\tilde{g}, \tilde{g} \rightarrow q\bar{q}\tilde{\chi}^0_1$. The gluino decay branching ratios are set to 50% for both the $dd\tilde{\chi}^0_1$ and $u\bar{u}\tilde{\chi}^0_1$ channels. Two values of the gluino mass are studied, $m_{\tilde{g}} = 1.4$ TeV and 2 TeV, with varying lifetimes between $\tau_{\tilde{g}} = 0.003$ ns and 50 ns. The neutralino mass is either set to $m_{\tilde{\chi}^0_1} = 100$ GeV or kept at $m_{\tilde{g}} - 100$ GeV. We generate events internally to CheckMATE, via the PyTHIA 8 interface, also carrying out the hadronization. Jets are reconstructed with FASTJET 3 \cite{49}, using the anti-kt algorithm \cite{50}, with a distance parameter of 0.4. The cross-section is varied by rescaling the number of events by the ratio (imposed cross-section)/($\text{SUSY}$ cross section), so that the limits are independent of the actual SUSY production cross-sections (or their LO calculation). Samples of $10^6$ events are considered for the relevant mass, lifetime and cross-section values.

While we principally consider unmatched events, we will also discuss the impact of jet radiation on individual scenarios. For this, we employ MADGRAPH 5 2.7.0 \cite{22} and add up to two jets using the MLM method with $x_{\text{cut}}$ parameter of 100 GeV.

Event selection

The event selection follows the recast instructions provided by the ATLAS collaboration \cite{33}. It includes a preselection defining the event- and vertex-level acceptances for truth-level particles, and then the application of parameterized efficiencies. This recast strategy has been successfully applied in the past (see Contribution 22 in \cite{36} and the publicly available codes \cite{34,35}).

Preselection

The missing energy $E_{\text{miss}}$, is defined at truth-level as the vector sum of the momenta of the stable invisible weakly-interacting particles ($\tilde{\chi}^0_1$ and neutrinos in the considered benchmark scenario). It is requested to satisfy the cut: $E_{\text{miss}} > 200$ GeV.

Jet requirements are applied on 75% of the events. These demand either one truth jet with $p_T > 70$ GeV or at least two jets with $p_T > 25$ GeV, satisfying in both cases a trackless requirement: the scalar sum of the charged particle $p_T$ in the jet should not exceed 5 GeV for particles with small impact parameter ($d_0$). We interpret the small impact-parameter condition as $d_0 < 2$ mm.

DV's are reconstructed from stable charged particles. The event should contain at least one DV in the fiducial volume: the transverse distance from the interaction point should lie between 4 mm and 300 mm, as well as $|z| < 300$ mm. The DVs should contain at least 5 selected decay products, i.e. stable and charged particles with $p_T > 1$ GeV and an approximate transverse impact parameter $d_0 > 2$ mm. The condition on the $p_T$ of the selected decay product assumes a charge $|q| = 1$, which we regard as implicitly fulfilled.

Efficiencies

Efficiencies of two origins are considered. The first one corrects $E_{\text{miss}}$. The second one applies on the reconstruction of the DV. Both are provided by the experimental collaboration \cite{51}.

The results presented in Fig. 2 seem to indicate a satisfactory performance of the recast strategy in scenarios involving a sizable mass-splitting between the LLP and the LSP. In addition, we were able to largely recover the cutflow information \cite{33} (as far as a comparison between truth- and detector-level events is possible).
The poor performance of the recast in this compressed configuration hints at the possible failure of the approach with parametrized efficiencies to simultaneously address all relevant spectra. As a possible way to mitigate incomplete efficiency parametrisations, the LHC Reinterpretation Forum white paper [52] and the LLP white paper [16] recommend using multiple topologies or multiple mass benchmarks for unweighting.

Matching events

Finally, we discuss the impact of jet radiation (matched events) for a gluino lifetime of 1 ns. The results are collected in Table 7, allowing for up to two radiated jets. We observe that the radiation of jets affects the limiting cross-section by only a few percent in the scenario with large gluino-neutralino mass-splitting, and by $\sim 10\%$ in the compressed case. Matching events thus leads to a minor effect in the considered search, justifying our focusing on unmatched events in Figs. 2 and 7. We note that the added jets affect the
limiting cross sections in opposite directions for the two scenarios, relaxing the bound in the uncompressed case, while tightening it in the other case. This can be qualitatively understood as, in the compressed case, the radiation of additional jets provides possibly more energetic candidates for the jet cuts, while the jets originating in gluino decays have reduced energy due to the narrow phase space. On the contrary, in the scenario with large mass-splitting, the jets originating in gluino decays are already energetic, so that radiating further jets decreases their ability to withstand the jet cuts.

Appendix D: Implementation details: displaced vertex and a muon

Monte Carlo samples

As a benchmark process for the search involving displaced vertices and displaced muons [29] the stop pair production was considered, followed by the RPV decay: \( pp \rightarrow \tilde{t}_1\tilde{t}_1, \tilde{t}_1 \rightarrow \mu q \). The Monte Carlo samples for cutflows and exclusion limits were generated using MADGRAPH 2.7.2 [22] with up to two additional partons matched using CKKW-L procedure [53]. Showering was performed by PYTHIA 8.244 [23]. To minimise the impact of statistical uncertainty the sample for each grid-point or cutflow scenario were 10–20 times larger than the nominal number of events. The samples were normalized to the approximate next-to-next-to-leading order in the strong coupling constant with a soft gluon resummation (approximate NNLO+NNLL) [54–57].

Event selection and signal regions

The event selection defines two mutually exclusive trigger-based signal regions: \( E_T^{\text{miss}} \) Trigger SR and Muon Trigger SR. The first SR should be selected by the MET trigger and requires \( E_T^{\text{miss}} > 180 \) GeV and a muon with \( p_T > 25 \) GeV, while the second one should be selected by the muon trigger with muon \( p_T > 60 \) GeV and \( E_T^{\text{miss}} < 180 \) GeV. For both signal regions the muon vertex should be at least 2 mm away from the primary vertex. The selection also requires a displaced vertex at least 4 mm away from the primary vertex, with at least 3 associated tracks and its visible invariant mass calculated from the four-momenta of the associated tracks \( m_{DV} > 20 \) GeV (assuming each track originates from a charged pion).

Displaced vertices are reconstructed internally by CheckMATE using a DVfinder class. Firstly, charged stable particles with production vertex separated from the primary vertex are searched for. These are then merged into vertices based on the relative position, if the distance is smaller than 1 mm. The position of the combined vertex is calculated as a weighted (with invariant momentum) average of the positions of input vertices. The procedure continues until no further vertices can be combined and the remaining single track vertices are dropped.

Validation

For the reconstruction of events at the CheckMATE level no specific truth particles efficiencies were used (these are available on HEPDATA). The ATLAS analysis rejects vertices from interaction with dense detector material. The veto is imposed based on a three dimensional map of the detector and rejects 42% of the events. Because the map was not public when the CheckMATE analysis was implemented, the veto is applied as a flat rejection probability on each reconstructed displaced vertex. Due to inhomogeneous distribution of vetoed points this can potentially result in a bias for certain lifetimes compared to the experimental analysis.

The validation procedure was performed both in terms of the exclusion contour for the benchmark model, see Fig. 3, and of the cutflows for three different parameter points. As already noted in Sect. 2.3, generally a good agreement is observed for the exclusion contour, except a range of lifetimes 0.01–0.1 ns where the recast exclusion is significantly weaker, though still within the 2-sigma band. In order to get a better insight, in Table 8 we also compare cutflows published by the ATLAS collaboration for \( m_{\tilde{t}} = 1.4 \) TeV and three lifetimes.

For the longest lifetime, 1 ns, the number of reconstructed events within CheckMATE is by 40% higher than that of ATLAS. This corresponds to the slightly stronger recast exclusion observed in Fig. 3 for lifetimes \( \sim 0.2–6 \) ns. For the intermediate lifetime, 0.1 ns, the number of CheckMATE reconstructed events is half of the ATLAS number. This corresponds to a weaker exclusion observed in Fig. 3 for lifetimes \( < 0.2 \) ns. We note however that this falls inside the expected 2-\( \sigma \) exclusion limit calculated by the collaboration. Finally, for the shortest lifetime, 0.01 ns, we observe almost a 3-fold difference. However, this region is close to the experimental sensitivity limit, as can be seen in Fig. 3, where the exclusion line becomes almost vertical. For this reason a poor agreement between the full simulation and recasting is not surprising. On the other hand, it also has a minimal impact on the exclusion contour.
Table 8 Cutflow for the displaced vertex + muon search in the missing transverse energy SR

| Selection | $\tau = 0.01$ ns | $\tau = 0.1$ ns | $\tau = 1$ ns |
|-----------|-----------------|----------------|--------------|
| ATLAS     | CM              | ATLAS          | CM           |
| All       | 64.2            | 64.2           | 64.2         |
| $E_T^{\text{miss}}$ trigger | 63.0 | 63.0 | 63.2 |
| $E_T^{\text{miss}} > 180$ GeV | 60.9 | 61.2 | 61.0 | 60.6 | 61.0 |
| $\geq 1\mu; p_T > 25$ GeV, $|\eta| < 2.5$ | 57.8 | 60.8 | 58.7 | 61.0 | 52.5 | 60.6 |
| $2 < |d_0(\mu)| < 300$ mm | 11.3 | 12.8 | 49.1 | 52.5 | 49.5 | 59.2 |
| $|z_0(\mu)| < 500$ mm | 11.3 | 12.8 | 49.1 | 52.5 | 49.3 | 58.1 |
| Fake/HF/cosmic veto | 9.1 | 9.9 | 40.0 | 42.4 | 39.4 | 48.1 |
| At least one DV | 8.5 | 4.4 | 37.6 | 29.8 | 32.6 | 39.6 |
| DV fiducial volume | 8.4 | 3.7 | 37.1 | 29.1 | 31.2 | 32.7 |
| Material veto | 5.3 | 2.2 | 31.0 | 16.9 | 22.2 | 19.0 |
| $n_{\text{tracks}} \geq 3$ | 3.8 | 1.8 | 26.0 | 15.5 | 13.7 | 17.3 |
| $m_{\text{DV}} > 20$ GeV | 3.4 | 1.2 | 22.7 | 11.9 | 10.3 | 14.0 |

Appendix E: Implementation details: heavy charged particle track

Monte Carlo simulation samples

In order to validate our implementation of this search in the CheckMATE we simulated the production of pairs of charginos and staus with MG5_AMC@NLO version 2.6.6 [22] including two additional partons at leading order in combination with PYTHIA 8 [23]. Matching was performed using the MLM scheme [58]. The resulting events in HepMC format [59] were processed by CheckMATE which uses DELPHES for the simulation of the ATLAS detector.

The reference benchmarks used for the validation are provided in SLHA format by the ATLAS collaboration in the HEPData repository [60]. The pair production of charginos is based on a mAMSB [61] scenario while the staus production on a GMSB [62] scenario. For reproducing the ATLAS results we have chosen a set of chargino and stau masses and simulated $2 \times 10^4$ events for each of the benchmarks considered.

Preselection

The following event preselection is applied:

- Full-detector candidates have to be electrically single-charged particles and pass the full ATLAS detector before decaying ($R > 12.0$ m, $|z| > 23.0$ m).
- Events with $E_T^{\text{miss}} > 300$ GeV are accepted, otherwise a trigger efficiency is applied.
- Single-muon trigger objects have to reach the ATLAS muon spectrometer (stable within $R > 12.0$ m, $|z| > 23.0$ m), have a $\beta > 0.2$, a pseudorapidity $|\eta| < 2.5$ and a minimum transverse momentum of 26 GeV. In addition a muon trigger efficiency which is a function of $\beta$ and $|\eta|$ of the particle is applied.

Maps of trigger efficiencies are provided by the ATLAS collaboration in the HEPDATA repository [63].

Search regions

The searches for pair-produced long-lived staus and charginos are based on a loose and on a tight candidate selection. The main difference between them is that the loose selection requires two candidates while the tight one only one. In both cases the candidates are required to have $p_T > 70$ GeV and while for the loose selection $p > 100$ GeV is imposed, the tight one requires $p > 200$ GeV. In addition a tighter pseudorapidity $|\eta| < 1.65$ for the tight selection is required. Finally trigger efficiencies for loose and tight candidates in function of $\beta$ and $|\eta|$ of the candidates are applied.

The signal regions are based on cuts on the loose and tight selections plus applying inclusive cuts on the time of flight (ToF) measurements of the particle mass $m_{\text{ToF}}$. The reconstructed $m_{\text{ToF}}$ is given as a function of the truth mass for the candidates and is sampled from a Gaussian using the mean and width from the respective bin in the truth mass. For candidates passing the loose cuts two masses have to be sampled for a given truth mass and the lower of the two is used for the final counting. For tight candidates only one

Table 9 Definition of the signal regions implemented in CheckMATE

| $m_{\text{ToF}}$ [GeV] |
|-----------------------|
| tight                |
| 175                   |
| 375                   |
| 600                   |
| 825                   |
| loose                |
| 150                   |
| 350                   |
| 575                   |
| 800                   |
mass is sampled using the respective mean and resolution. The tables parametrizing the reconstructed $m_{\text{TOP}}$ are provided in the HEPDATA repository [63].

Following the prescription above, the signal regions are defined according to the loose and tight cuts plus a condition on the minimum reconstructed $m_{\text{TOP}}$ as summarized in Table 9. The results of the validation are presented in Fig. 4 and establish the reliability of this recast search in CheckMATE.

Appendix F: Implementation details: disappearing track

Monte Carlo simulation samples

We generated the SUSY signal samples in the framework of minimal AMSB model [64] assuming $\tan \beta = 5$, with a positive higgsino mixing mass parameter, while the scalar masses are decoupled by setting $m_0 = 5$ TeV. The chosen benchmark contains a chargino with a 0.2 ns lifetime and a mass of 400 GeV. For the strong production channel, we select a gluino mass of 1600 GeV, a chargino mass of 500 GeV and of 400 GeV. For the strong production channel, we select a gluino mass of 1600 GeV, a chargino mass of 500 GeV and 0.2 ns lifetime and a mass of 400 GeV. The SUSY mass spectrum files are generated with ISASUSY 7.80 [65]. The gluino decays via off-shell squarks into the following decay channels,

$$
\tilde{g} \rightarrow q\bar{q}'\tilde{\chi}^0_1, \; \bar{q}q'\tilde{\chi}^-_1, \; \bar{q}q'\tilde{\chi}^+_1. \tag{B.2}
$$

assuming only first and second generation partons intervene. The wino–like chargino LLP decays into a charged pion and a neutralino LSP.

The MC events for the signal are generated together with two additional radiated partons in the hard process with MADGRAPH5 2.6.1 [22]. The parton level events were showered and hadronized with PYTHIA 8.230 [23]. We used the NNPDF 2.3LO parton distribution function [66]. Renormalisation and factorisation scales were kept at the default scale of MADGRAPH5. The combination of the parton shower and the matrix elements partons was performed in the CKKW-L merging scheme [67] and the merging scale was set to a quarter of the wino mass for the electroweak production channel or a quarter of the gluino mass for the strong production channel. The cross sections for electroweak production channel are calculated at NLO using PROSPINO2 [68] and the normalisation for the strong production is determined at NLO+NLL accuracy with NLLFAST [69–73].

Preselection

We apply the object removal procedure given in Ref. [74]. Electron and muon candidates must satisfy $p_T > 10$ GeV and $|\eta| < 2.47$ and 2.7, respectively. The isolation criteria for final state leptons are given by the requirement that the scalar sum of the transverse momentum of tracks inside variable-size cone around leptons is less than 15% of lepton transverse momentum while the cone radius is defined as min($\Delta R = 10$ GeV/$p_T$, $\Delta R = 0.2$ (0.3)) for electrons (muons). Jets are reconstructed with anti-$k_T$ algorithm and $\Delta R = 0.4$. Jet candidates must pass $p_T > 20$ GeV and $|\eta| < 2.4$. The missing transverse momentum is given as the negative vector sum of all reconstructed detector level objects, i.e. electrons, muons, photons and jets.

Search regions

For the electroweak and strong signal region, a set of common kinematic preselection cuts are demanded and for both signal regions we require a lepton veto, i.e. no isolated electrons and muons are allowed in the event. Two signal regions targeting the electroweak production channel (EW SR) and strong production channel via gluinos (strong SR) are considered. We apply parton level cuts on the tracklet (i.e. the long lived chargino). We demand isolated charginos with $p_T > 100$ GeV. Moreover, we require geometric acceptance cuts $0.1 < |\eta| < 1.9$. For the EW SR, the events must have at least one jet with $p_T > 140$ GeV and $E_T^{\text{miss}} > 140$ GeV (90 GeV < $E_T^{\text{miss}}$ < 140 GeV) in the high- (low-) missing transverse momentum region. Multijet background suppression is achieved by a $\Delta \Phi$ cut. Here, the difference in azimuthal angle between the missing transverse momentum and each of the up to four highest-$p_T$ jets with $p_T > 50$ GeV is required to be larger than 1.0.

The strong SR demands that events have a jet with $p_T > 100$ GeV, at least two additional jets with $p_T > 50$ GeV and $E_T^{\text{miss}} > 150$ GeV (100 GeV < $E_T^{\text{miss}}$ < 150 GeV) in the high-(low-) missing transverse momentum. The $\Delta \Phi$ between the missing transverse momentum and each of the up to four leading jets with $p_T > 50$ GeV is required to be larger than 0.4.

References

1. ATLAS Collaboration, Search for squarks and gluinos in final states with jets and missing transverse momentum using 139 fb$^{-1}$ of $\sqrt{s} = 13$ TeV $pp$ collision data with the ATLAS detector. JHEP 02, 143 (2021). https://doi.org/10.1007/JHEP02(2021)143. arXiv:2010.14293
2. M. Ibe, T. Moroi, T.T. Yanagida, Possible signals of wino LSP at the large hadron collider. Phys. Lett. B 644, 355 (2007). https://doi.org/10.1016/j.physletb.2006.11.061. arXiv:hep-ph/0610277
3. M. Cirelli, N. Fornengo, A. Strumia, Minimal dark matter. Nucl. Phys. B 753, 178 (2006). https://doi.org/10.1016/j.nuclphysb.2006.07.012. arXiv:hep-ph/0512090
4. P. Fileviez Perez, H.H. Patel, M.J. Ramsey-Musolf, K. Wang, Triplet scalars and dark matter at the LHC. Phys. Rev. D 79, 055024 (2009). https://doi.org/10.1103/PhysRevD.79.055024. arXiv:0811.3957
